R E P O R T R E S U M E S

ED 018 835 CG 001 029

THE NATIONAL APTITUDE SURVEY. FORMAL REPORT AND TEST MANUAL.

BY- KALTSOUNIS, GEORGE L. AND OTHERS
MICHIGAN UNIV., ANN ARBOR,SCH. OF EDUCATION
AUTOMOBILE MANUFACTURERS ASSN., DETROIT, MICH.

PUB DATE FEB 68

EDRS PRICE MF-$0.50 HC-$2.96 72P.

DESCRIPTORS- *APTITUDE TESTS, *OCCUPATIONAL GUIDANCE, *HIGH SCHOOL STUDENTS, *PREDICTION, LONGITUDINAL STUDIES, VOCATIONAL COUNSELING, TEST CONSTRUCTION, OCCUPATIONAL INFORMATION, PREDICTIVE VALIDITY, FOLLOWUP STUDIES, SOCIOECONOMIC BACKGROUND, *AUTO MECHANICS,

THE NATIONAL APTITUDE SURVEY ATTEMPTED TO PROVIDE A HIGH SCHOOL GUIDANCE INSTRUMENT TO IDENTIFY POTENTIAL AUTOMOTIVE REPAIRMEN. THE FORMAL REPORT INTRODUCES A GENERAL STATEMENT WITH THE NEEDS, THE ADVANTAGES, THE DEVELOPMENT, AND THE SAMPLE FOR THE SURVEY. THE SURVEY TEST MANUAL DISCUSSES ADMINISTRATION AND SCORING, NORMS AND PROFILES, RELIABILITY, VALIDITY, VALIDATED TEST USE, AND A SUMMARY. THE SURVEY DEVELOPED THROUGH FOUR STAGES--(1) THE TEST BATTERY'S PRESENT FORM, (2) SAMPLE PREPARATION AND ACTUAL ADMINISTRATION, (3) CLERICAL MAINTENANCE OF PARTICIPANT CONTACT THROUGH THREE FOLLOW-UPS IN SIX YEARS, AND (4) BATTERY VALIDATION AND REPORT COMPILATION. THE SIX-TEST BATTERY MEASURES FACTORS CHARACTERIZING THE OCCUPATIONAL BEHAVIOR OF EMPLOYED AUTOMOTIVE REPAIRMEN. A BIOGRAPHICAL INVENTORY INVESTIGATES SOCIO-ECONOMIC AND ATTITUDINAL FACTORS. A RANDOMLY SELECTED, NATIONALLY REPRESENTATIVE PUBLIC HIGH SCHOOL SAMPLE OF 7,265 NINTH AND 10TH GRADE BOYS WERE IDENTIFIED AND TESTED IN 1961, WITH COMPLETED INFORMATION EVENTUALLY OBTAINED FOR 80 PERCENT. THE AUTOMOTIVE AND GENERAL VOCABULARY TESTS AND THE SIX BIOGRAPHICAL INVENTORY ITEMS WERE FOUND TO BE THE BEST PREDICTORS. THE RESULTING GUIDE ESTIMATES A BOY'S CHANCES OF BECOMING A REPAIRMAN WITHIN THREE YEARS OF GRADUATION. (WR)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE UNIVERSITY OF MICHIGAN

SCHOOL OF EDUCATION

ED018835

The National Aptitude Survey

Formal Report and Test Manual

GEORGE L. KALTSOUNIS
ROBERT J. CROWLEY
DAN N. PERKUCHIN

CG 001 829

Sponsored by:

The University of Michigan
and the
Automobile Manufacturers Association

Administered through:

February 1968

OFFICE OF RESEARCH ADMINISTRATION • ANN ARBOR

THE NATIONAL APTITUDE SURVEY

Formal Report and Test Manual

George L. Kaltsounis
Professor of Education
State University of New York—College at Buffalo

Robert J. Crowley
Assistant Professor of Education
Colgate University, Hamilton, New York

Dan N. Perkuchin
Research Assistant
The University of Michigan

Ralph C. Wenrich, Project Director

The University of Michigan
and the
Automobile Manufacturers Association

February 1968

The individuals who worked with him on the
National Aptitude Survey
dedicate this report to

HUGH F. PIERCE
Project Director (1957-1964)

The University of Michigan—Automobile Manufacturers Association
Personnel Classification Project

ADVISORY COMMITTEE AND STAFF
of
The University of Michigan—Automobile Manufacturers Association Project

Members of the Advisory Committee

Representatives of the University of Michigan

Dr. M. Clemens Johnson	Professor of Education
Dr. John E. Milholland	Professor of Psychology
Dr. Edward C. Roeber	Professor of Education
Dr. Garry R. Walz	Associate Professor of Education
Dr. Frank B. Womer	Associate Professor of Education
ex officio	
Dr. Willard C. Olson	Dean of the School of Education Professor of Education Professor of Psychology
Mr. Ralph C. Wenrich	Chairman, Department of Vocational Education and Practical Arts

Representatives of the Automobile Manufacturers Association

Mr. E. C. Henricks	Service Managers Committee
Dr. Robert C. Lusk	Secretary, Service Managers Committee
Mr. Stanley S. Roe	Formerly Secretary, Service Managers Committee
Mr. Harry A. Williams	Managing Director

Representatives of the Michigan Department of Public Instruction

Mr. James H. Dotseth	Formerly, Consultant, Trade and Industrial Education Division of Vocational Education
Mr. John J. Harris	Assistant Superintendent Trade and Industrial Education Division of Vocational Education
Mr. Robert M. Winger	Formerly, Assistant Superintendent Division of Vocational Education

Resource Persons

Mr. A. A. Althouse	Detroit Public Schools
Mr. C. T. Doman	Formerly, National Service Manager Ford Motor Company
Miss B. Dvorak	U. S. Bureau of Employment Security
Mr. P. B. Hopkins	Chrysler Corporation
Mr. W. Landon	Chrysler Corporation
Mr. L. A. Mitchell, Jr.	General Motors
Mr. B. R. Sweeney	Ford Motor Company
Mr. M. E. St. Aubin	General Motors
Mr. C. M. Tillinghast	American Motors

STAFF

Project Directors

Hugh F. Pierce (1957-1964)	Late Lecturer, Department of Vocational Education and Practical Arts
Ralph C. Wenrich (Acting Director) (1965-1967)	Professor and Chairman, Department of Vocational Education and Practical Arts
Almando A. Vezzani (Acting Director) (Jan.-Apr. 1966)	Late Professor of Vocational Education and Practical Arts

Research Associates

Dr. Joseph W. Dorton
Dr. George L. Kaltsounis
Dr. Robert J. Crowley

Research Assistants

Mrs. Allison C. Burleigh
Mr. Jason Millman
Mr. Dan N. Perkuchin

ACKNOWLEDGMENTS

The National Aptitude Survey represents a decade of effort expended to provide American high schools with a guidance instrument especially designed for the identification of potential automotive repairmen An essential element of the basic research design was longitudinal prediction, which consisted of a six-year follow-up of ninth and tenth grade high school students identified and tested in 1961. The Formal Report relates the stages in the developmental history of the design's implementation.

The roster of individuals associated with the project includes: Professor Ralph C. Wenrich, who succeeded Mr. Hugh Pierce as Project Director; Drs. Joseph W. Dorton and George L. Kaltsounis, who conceived, developed and standardized the test battery; Dr. Robert J. Crowley, the editor of this report, who in conjunction with Mr. Dan N. Perkuchin carried out the final stages of criteria validation; and Mrs. Allison C. Burleigh and Mr. Jason Millman, who helped in the test development stage.

Mrs. Helen F. MacDonald and Mrs. Ann M. Rynning maintained the clerical and secretarial records over a span of years; Mrs. Hollis Westerman managed the details of the last mailings; and Miss Beth Rosengard and Miss Jeannine Horton typed the final manuscript.

Mr. Ray Bollman of Chrysler Motors, Mr. Lewis C. Harris and Mr. Paul McDonald of General Motors, and Mr. T. K. Rossiter of Ford Motor Company provided advice and assistance when asked. The necessarily anonymous service managers in dealerships, local and national, complied wholeheartedly with the many requests made of them. Mr. Gilber L. Haley of the Michigan Automobile Dealers Association introduced the research staff to the members of his organization. Especially kind and helpful were service officials at American Motors, Chrysler, Ford and General Motors.

Doctors M. Clemens Johnson, John E. Milholland, Edward C. Roeber, Garry R. Walz, and Frank B. Womer, who constituted the University Advisory Committee, devoted hours of their time, singly or as a committee, pondering the successive steps of the development and making specific recommendations to implement them.

Dr. Willard C. Olson, Dean of the School of Education, worked closely with the project directors in administrative matters and provided far more than the necessary counsel and support such a long-range undertaking requires.

Dr. Robert C. Lusk, who acted as liaison between the University and the Automobile Manufacturers Association, co-sponsors of the project, exercised

his considerable tact, broad understanding, and congenial wit to facilitate the progress of the project.

Ralph C. Wenrich
Acting Project Director

TABLE OF CONTENTS

		Page
LIST OF TABLES		xi
LIST OF FIGURES		xii
	A. Formal Report	
I.	INTRODUCTION	3
	General Statement about the National Aptitude Survey	3
	The Need for the National Aptitude Survey	3
	The Advantages of the National Aptitude Survey	6
	The Development of the National Aptitude Survey	6
	The Sample	7
II.	THE TESTS AND THE PROBLEM	11
	The Tests of the National Aptitude Survey	11
	B. Test Manual	
III.	ADMINISTRATION AND SCORING	19
	General Statement	19
	The Materials Needed for the Test	19
	Suggestions for Scheduling Testing Programs	20
	General Considerations for Administration	20
	Steps in Administering the Tests	21
IV.	NORMS AND PROFILES	24
	General Statement	24
	Reading the Score-Distribution Tables	24
	Reading the Norm Tables	24
	Intercorrelations of the Tests	30
V.	RELIABILITY	32
	General Statement	32
	Types of Reliability Coefficient Used	32

TABLE OF CONTENTS (Concluded)

VI.	VALIDITY	35
	General Statement	35
	Validational Sample	35
	Analysis of Test Scores of the Validational Samples	37
	Summary	43
VII.	USE OF THE VALIDATED TESTS	45
VIII.	SUMMARY AND RECOMMENDATIONS	47
APPENDIX A.	NATIONAL APTITUDE TEST SCORES AS PREDICTORS OF ACHIEVEMENT IN HIGH SCHOOL VOCATIONAL EDUCATION COURSES (Kaltsounis' Study)	49
APPENDIX B.	BIOGRAPHICAL INVENTORY ITEMS SELECTED BY THE REGRESSION PROGRAM	53
BIBLIOGRAPHY		55

LIST OF TABLES

Table		Page
1	A Comparison of the Normative Samples by Respondent Category	10
2	Time Limits for Administration of the National Aptitude Survey	22
3	Frequency Distribution of Raw Scores on Tests of the National Aptitude Survey (Probability Sample)	25
4	Frequency Distribution of Raw Scores on Tests of the National Aptitude Survey (Non-Probability Sample)	26
5	Intercorrelations of All Tests of the National Aptitude Survey for the Entire Study Sample (N=1032) and the Auto Shop Group (N=31) (Kaltsounis' Study)	31
6	Estimated Reliabilities and Standard Errors of Measurement for the National Aptitude Survey (Probability Sample)	32
7	Coefficients of Reliability and Other Statistics on the National Aptitude Survey for the Study Sample and One Coefficient for the National Sample	33
8	Distribution of Participants into Relevant Subsamples	36
9	Multiple Correlation Coefficients and Coefficients of Determination for Eight Selected Variables (322 MECSAM vs. 907 SUBSAM)	38
10	Percentile-Range-Composite-Score Boundaries, Proportions of Automotive Repairmen, and Likelihood Estimates of Becoming an Automotive Repairman	40
11	Predictors and Criteria (Kaltsounis' Study)	50
12	Validity Coefficients of the National Aptitude Survey for Twelve Criterion Groups (Kaltsounis' Study)	51
13	Regression Coefficients, Coefficients of Correlation and Determination, and Standard Errors of Estimate for the Corresponding Criterion Groups (Kaltsounis' Study)	52

LIST OF FIGURES

Figure		Page
1	The Individual Profile Form (Probability Sample)	27
2	The Individual Profile Form (Non-Probability Sample)	28
3	A Profile Form for an Hypothetical Individual (Probability Sample)	29
4	A Comparison of Percentages obtained with Two Combinations of Variables	41
5	A Comparison of Percentages of Automotive Repairmen within Percentile Ranges for the Validational Group (N = 6328) and a Cross-Validational Group (N = 1113)	44

A. FORMAL REPORT

I. INTRODUCTION

GENERAL STATEMENT ABOUT THE NATIONAL APTITUDE SURVEY

The purpose of this battery of tests, entitled the National Aptitude Survey, is to identify future automotive repairmen. It provides the basis for a prediction as to how likely it is that a male student in the ninth or tenth grade of a public high school, who has attained a particular score, will enter the field of automotive repairmen within three years after graduation. Those students identified as being the likeliest future automotive repairmen, consequently, could benefit the most from school courses oriented to such a career. A boy's chances of being an automotive repairman is determined by comparison of his obtained score for the appropriate tests with a normative distribution of scores obtained from nearly all the participants in the study. The prediction of a boy's chances takes the form of a likelihood estimate which indicates the increase in accuracy to be gained from knowing a test score in contrast to not knowing it.

The National Aptitude Survey incorporates three important aims: specificity in its orientation to a particular field; representativeness in its normative group; and longitude in its predictions. It is specific in that it is intended to predict entrance and success in the occupation of automotive repairman; representative in that its normative population of ninth-grade males represents the general ninth grade students for whom the Survey is designed; and longitudinal in that test performance at the ninth and tenth grade levels serves to predict for each student the probabilities of successful performance as a mechanic several years after testing. These features should make this instrument especially useful to guidance counselors in their counseling of students at these same levels of schooling.

THE NEED FOR THE NATIONAL APTITUDE SURVEY

As the products of a technological society increase in number and attractiveness, concurrently there develops a need for a group of individuals skillful enough to service them. One ever-present danger has been a widening discrepancy between dispersion of a product and its maintenance. Consumer satisfaction with a purchase influences his attitudes toward newer developments in the same product.

In the automotive field, as in other areas, the quality and number of individuals who become automotive repairmen exert a strong influence on the customer's attitudes toward performance of the product. For some time, the percentage of qualified individuals entering the trade of repairman has not been sufficient to satisfy the demand. Changes in occupation, subsequent to their

employment, have further intensified the need. Yet, large numbers of unskilled youth are entering the available work force. That there exists a need for skillful automotive repairmen may be largely unknown or uninteresting to them at this stage of their career.

After high school, the range of occupational opportunities does permit a short period of trial-and-error choices, but at the expense of possible income and available time. Testing and counseling at this time under the direction of agencies within society may alleviate the situation somewhat by restricting the range of suitable and possible alternatives. Nevertheless, one could argue that counselors in high school could have guided those youth who showed an aptitude for automotive repair into courses which would have provided training prior to entry into the trade.

The University of Michigan was approached by Andrew Althouse, a representative of the AI-VE Conference (Automotive Industry-Vocational Education Conference) to consider the possibility of conducting research in the selection of auto mechanics. A proposal was prepared and submitted to the Service Manager's Committee of the Automobile Manufacturers' Association, and in September 1957 a grant of $140,000 was made by that association to The University.

At the time of the conceptualization of the National Aptitude Survey, there was no known instrument in existence designed specifically to predict success in the trade of automotive repairman which did not possess intrinsic limitations. A survey of the available instruments revealed features in their validation which militated against their use. Indications of low utility to the study were such features as lack of independent cross-validation, validation based upon individuals already employed, the limited size of the validational samples, limitations of the population to workers in a single plant or company, inflated correlation coefficients, and lack of published literature that evaluated the tests adequately. The National Aptitude Survey incorporated into its design features intended to circumvent the difficulties that plague other tests. These included:

a. validation based on a sample of participants who became automotive repairmen several years after the testing;

b. utilization of validational devices that would minimize error;

c. cross-validation upon a random subsample of the test participants.

The psychological assumption that underlies the National Aptitude Survey is that it is possible to specify some predictors of success in the career field of automotive repairmen. The identification of possible factors which could contribute to an aptitude in this career field was the focus of the initial efforts in developing the test battery. The subsequent pilot investigation encompassed a survey of related research and the compilation of detailed "job analyses."

An analysis* of the records made available by the Connecticut vocational schools provided information about the predictive power of some of the factors considered relevent. Knowledge of the Abstract Reasoning scores obtained through administration of the Differential Aptitude Tests improved prediction of both Auto Theory grade and overall grades in technical training by 29% and 23% respectively (N = 22) over no such knowledge at all. This same set of scores, however, was not predictive of occupation after graduation (N = 18). Number of children in the family proved useful in predicting entrance into the field of automotive mechanics, an improvement of 24% over ignorance of such information.

A more direct source of information was an intensive series of job analyses by the first of the co-authors (63). These analyses, combined with a thorough review of the literature, provided a list of factors of abilities (such as intelligence, space perception, and reasoning), factors of past learning (previous job experience and type and length of training), personality factors (temperament, emotional stability, rigidity, and patience), as well as factors of interest (hobbies, values, and goals) and of a biographical nature (residence, number of children in the family, occupation of parents), all of which are related to entrance and success in the automotive field.

Perceptive individuals have long recognized the need for specific aptitude tests in place of, or as a supplement to, tests of general mental ability. In the last three decades, a number of such tests have appeared, and the National Aptitude Survey represents a contribution to this trend.

The definition of aptitude utilized by the Differential Aptitude Tests (9, p.2) reflects accurately the thinking which guided construction of the present battery:

> Aptitude: A condition or set of characteristics regarded as symptomatic of an individual's ability to acquire with training (usually specified) some knowledge, skill, or set of responses, such as the ability to speak a language, to produce music . . .

The National Aptitude Survey avoids the assumption that the special aptitude sought is an innate characteristic of the individual acquired through heredity. Inheritance may limit the potentialities of the individual, but within limitations a wide variety of learnings is possible. The Survey demonstrates that the individual performs in a certain way which is related to the behavior of automotive repairmen in a work situation. The demonstration of this relationship involves a comparison between the behavior shown by this

*"Analysis of Existing School Records from Vocational Schools," The University of Michigan—Automobile Manufacturers Association Research Project, 1960 (unpublished study).

same individual who has become a repairman, and the behaviors he displayed in the test situation several years earlier. The term, "behavior," in this context represents an interaction of his heredity and environment, the result not only of his potential, but also of his learning. His intelligence, achievement, personality interests, and skills, as well as the nature and quality of the opportunities inherent in his environment, all contribute to the measured behavior. The expectancy of success, derived from valid predictors, obviates the need to specify as either innate or acquired the origin of any characteristic.

THE ADVANTAGES OF THE NATIONAL APTITUDE SURVEY

The possibility that more than one factor of ability contributed to success as an automotive repairman suggested that a combination of certain tests would be the best predictor. Each test would have to make its unique contribution, and failure to do so would result in its elimination.

Whenever there has existed a need for a combination of tests to measure an aptitude, one practice has been to assemble several worthwhile, but unique, tests to measure it. The very urgency of the need frequently has led to toleration of what is basically a questionable practice. Each such test presumably underwent its own independent development and the subsequent linkage of disparate tests has usually left unanswered the question as to the nature of the new relationships between them. A gratuitous assumption of equivalence of population and purposes should not be considered merely incidental to the interpretation of results. The development of the National Aptitude Survey proceeded with one population and purpose in mind, at almost the same period of time; the interpretation of the relationships between its tests requires no assumption of equivalence for the following reasons:

a. The fact that all subtests were standardized at the ninth- and tenth-grade levels, the levels at which it is to be used, should eliminate the difficulties that arise from using a combination of tests standardized on population other than the desired one;

b. Whereas a use other than the intended one may result in differences in matters such as format, administrative procedures, the use of answer sheets, and time limits other than those originally specified, the National Aptitude Survey tests avoid these liabilities due to the care and industry that in each instance attended their experimental determination on a specified population.

THE DEVELOPMENT OF THE NATIONAL APTITUDE SURVEY

The final form of the battery is the result of an effort to subject each item to pretesting, statistical analysis and revision. The testing occurred

in three phases: pretesting (ninth-grade students in the Ann Arbor Schools); local testing (schools in the vicinity of Ann Arbor); and area testing (schools in southeastern Michigan and a northern section of Ohio). By the termination of this segment of the development phase the research team had eliminated several tests for reasons such as lack of reliability and/or lack of unique measurement, and had revised the length of others.

Two criteria guided the decisions to include an item in a particular test:

a. There had to be a high biserial correlation coefficient between an item's score and the total score on the test. The biserial correlation coefficients served as a measure of the internal consistency of the test.

b. Each item had to have a difficulty index which ranged from .15 to .85. A value on this index referred to the percentage of individuals that passed the item out of all who attempted it. The items in each test were representative of the entire span of allowable difficulties.

THE SAMPLE

Organization Of The Sample

The drawing of the sample followed the two requirements of a normative sample: it was large and representative. The sample size sought was 10,000, or one out of every 130 male students estimated from official publications to have been enrolled in ninth grades nationally, at the time prior to the national testing. The sample design specified a stratified, multi-stage cluster sample with probability proportional to size.

One feature of the sampling procedure was the use of a "measure of size." The design required that every pertinent school in the population of eligible schools located throughout the continental United States had a chance of being represented. There were 1548 schools in the 66 sampling Strata, denoted by the 66 Primary Sampling Units (PSU). The unit for the "measure of size" was 25. Schools with a 25±12 ninth-grade students received a "measure of size" 1; those with 50 ± 12, a size of 2; those with 75 ± 12, a size of 3, until all the schools had some numerical equivalent of size.

Another feature was a recombination, for administrative reasons, of all sampling units into clusters of fifty students. Schools with the "measure of size" 1 were combined with other similar schools to form a school of size 2; if selected, all the ninth-grade boys in these schools would take the tests. For schools with a "systematic" subselection was obtained. The formula for this procedure was $2/S_i$; in which S_i equalled the measure of the size of the ith school. Thus in a school of size

12, (N of ninth graders = 300 ± 12) the subselection ratio was 2/12 = 1/6— or one out of every six boys. By means of alphabetical lists of all ninth grade boys in the school it was possible to select those to be tested. Within the first six names at the top of the list the research staff selected a starting number at random, then by means of a constant interval of six chose the remaining names. A school of size 12 would yield a group of approximately 50 pupils for testing (1/6 of 300 ± 12 boys). The actual size of the group, however, depended on the accuracy of the estimate of the school, and on the proximity of the actual ninth-grade male enrollment to the midpoint of the class interval of "measure of size." Of two schools with the same "measure of size," e.g., 12, one could have yielded 48 boys (N=288), the other 52 (N=312).

Selection Of The Sample

The actual selection of the desired sample occurred during the summer of 1960. The resulting sample contained 236 schools out of a possible 1548, with 189 clusters of fifty boys out of a possible 4012 in the 66 P.S.U. Eighty-four of the 236 schools originally had a "measure of size" of less than 2; the remainder ranged from 2 to 32, with a median at 5.

Results Of The Selection And Subsequent Testing

Testing in the schools began in the Spring of 1961 and ended in the Fall of the same year. Out of 236 schools selected, 174 (74%) participated in the testing; the remainder did not find it convenient to take part in the program.

The tested sample corresponds very closely to the one selected in terms of the "measure of size" of the schools. There are no over-representations of certain size of schools or under-representations of others; the median representation of each level is 73% — that is, on the average about 73% of each size of a selected school is represented in the tested sample.

The Entire Sample And The Validation Sample

For the general male population, the rate of entry into the trade of automotive repairman is estimated to be about 2% yearly. It would require the selection of a sample size of nearly 10,000 to obtain a validational group of 200 cases. The establishment and determination of norms, regression equations, and expectancy tables made such a sample size imperative. The deficiency in numbers, which resulted from the less than total participation, obligated the research staff to augment the number of students tested, if an adequate sample of automotive repairmen was to be obtained. A second sample of 804 tenth grade boys in 12 Michigan and 9 Connecticut trade schools also took the tests. The percentage of students from such schools who enter the trade of automotive

repairman is considered to be higher than the 2% from the general population. This group is known as the Non-Probability Sample of Tenth-Grade Boys.

Longitudinal Aspects Of The Sample

During the year that followed each administration of the test battery, the Project staff began a procedure designed to maintain contact with every individual who had participated. Every year, at intervals of approximately fourteen months, each tested student, now known as a participant, received a questionnaire designed to elicit information about his school. He usually gave his present status, and any changes of address. After graduation, another questionnaire requested his work history. Those who became automotive repairmen became the validation sample. Information about the validation sample appears in the validation section(Section VI). There were three follow-up questionnaires in all.

By January, 1967, more than three years after the end of the First Follow-up, the rate of failure to maintain contact had increased (Table 1). The division of the table according to kind of sample obscures the fact that the time of the contacts was different. Each follow-up usually lasted more than fourteen months; the Survey staff would first attempt to contact the non-probability and a greater part of the probability sample. Mailings to the remainder of the probability sample followed seven to eight months later.* Yet the percentages of returns or nonreturns for either sample are similar.

---

*This procedure reflects the order of testing. One hundred seven schools, including the majority of those in the probability sample and the entire non-probability sample administered the tests in Spring, 1961. The remaining schools in the probability sample held testing sessions in Fall, 1961. The earlier schools would have graduated their students in Spring, 1964, three years after the test; the later group would have been graduated in Spring, 1965. At the time of the validational study each participant could have had at least two, and possibly three, years of experience as a mechanic after graduation.

TABLE 1

A COMPARISON OF THE NORMATIVE SAMPLES BY RESPONDENT CATEGORY

Normative Sample	Respondent Category	Sample Size as of:	N	Percentage of Potential Respondents	Remarks
Probability	Potential	A. Jan. 1, 1964	6820	--	Deceased: 44
		B. Jan. 1, 1967	6776	--	
	Actual	A. Jan. 1, 1964	6641	97.4	
		B. Jan. 1, 1967	5405	79.8	
	Lost	A. Jan. 1, 1964	91	1.3	
		B. Jan. 1, 1967	477	7.0	
	Non-response	A. Jan. 1, 1964	88	1.3	
		B. Jan. 1, 1967	894	13.2	
Non-probability	Potential	A. Jan. 1, 1964	804	--	Deceased: 10
		B. Jan. 1, 1967	794	--	
	Actual	A. Jan. 1, 1964	768	95.5	
		B. Jan. 1, 1967	631	79.5	
	Lost	A. Jan. 1, 1964	13	1.6	
		B. Jan. 1, 1967	51	6.4	
	Non-response	A. Jan. 1, 1964	23	2.9	
		B. Jan. 1, 1967	112	14.1	

II. THE TESTS AND THE PROBLEM

THE TESTS OF THE NATIONAL APTITUDE SURVEY

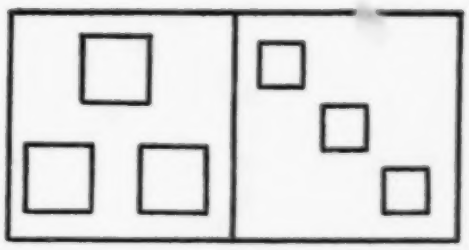
The National Aptitude Survey is a battery of six paper-and-pencil tests and a Biographical Information Blank designed for use at the junior-senior high school level. They are the following—each described briefly.*

Reasoning Test

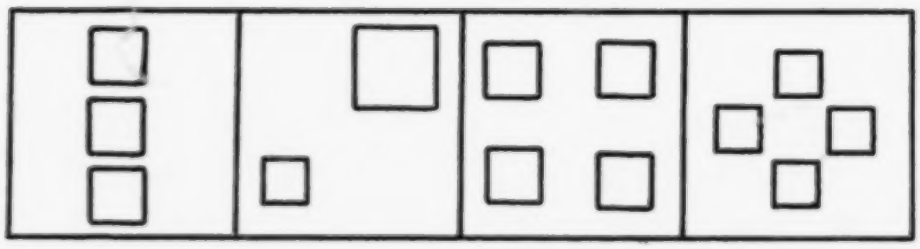
Except for the simple directions, this is a nonverbal test designed to be a measure of abstract nonverbal reasoning which is so eminently present especially in some technical vocations which determine inductively the logical relationship of underlying causes and observable abnormal symptoms of machines and equipment. It is, of course—and it has normally been (41)—an aspect of intelligence especially appropriate for those individuals who are considered to be culturally handicapped in language development.

The subject, in this test, is asked first to identify the concept common to two problem-figures and then select, from among four alternatives, the answer-figure that shares the same concept. A sample item follows.

PROBLEM BOXES



ANSWER BOXES



A+** B C D

The test has a total of thirty items, and it takes twenty-three minutes to administer.

---

*The tests are not yet commercially available. Single copies, however, may be obtained upon request from the Project Director, The University of Michigan—Automobile Manufacturers Association, School of Education, Ann Arbor, Michigan.

**Correct alternatives in sample items are indicated by plus signs.

General Vocabulary Test

Since vocabulary has traditionally played a prominent part in intelligence tests from the beginning of the mental testing movement (41, 66), and has been shown to correlate highly with overall measures of intelligence (98), this test was designed to be a measure of general verbal intelligence. Vocabulary tests of this type have been found to correlate with total measures of general intelligence as closely as do two different tests of general intelligence. (20, 41, 94, 98)

The subject, in this test, is asked to identify from among five alternatives, the vocabulary term that best fits the given definition. Two sample items follow.

The outside covering of a tree:

(1) root (2) cambium (3) blanket (4) bark+ (5) chlorophyl

A person temporarily residing in a place:

(1) transient+ (2) emigrant (3) settler (4) citizen (5) imigrant

Automotive Vocabulary Test

Discussing the Information Test of Project Talent, Dailey and Shaycoft state:

> It is hypothesized that the extent to which a person is interested in a specific area may be deduced from the shape of his profile as expressed by test results, and particularly from the relation between his level of information in the area in question and his level of information in other areas. <u>This hypothesis is based on the premise that the individual will tend to know most in the areas where his interest is greatest</u>* (20, p. 10).

This test was designed to be a measure of amount of knowledge of automotive terms, and indirectly, a measure of interest and experience in the automotive field. It is expected that students who score high on this test will tend to have high interest in this field and do well in training in automotive courses.

The structure of this test is identical with that of the General Vocabulary. The subject is asked to identify, from among five alternatives, the

*Emphasis is first author's.

technical term that best fits the given definition. A sample item follows.

The pin that attaches the piston to the connecting rod:

(1) wrist+ (2) cotter (3) drift (4) clevis (5) crank

The two vocabulary tests are combined in a single booklet, under one set of directions with the items alternating from general (odd) to automotive (even) vocabulary. There are sixty items in the entire booklet, thirty per each test, which takes twenty-three minutes to administer. The two tests were combined in one booklet under the assumption that the alternation of item content will reduce fatigue on the part of the examinee, and increase his interest for the test.

Chart Plotting Test

This is a new test in the field of aptitude measurement. It was designed to be a measure of the ability to follow simple written instructions and the extent of carefulness one applies in carrying them out. Success in a number of technical occupations in which the individual relies on manuals and other step-by-step instructions for the performance of his task is expected to depend to a high degree on the combination of factors measured by this test. Consideration has also been given to the possibility of obtaining, through this test, some indications of "the test-taking attitude" examinees assume at the time of the testing. Some preliminary work with this test has pointed out that students who take the testing seriously—that is, students whose "formula score" adjusted for random guessing on some of the other tests is very close to their "number right score"—tend to make higher scores on this test than those students who operate at a chance level throughout the range of item difficulty—that is, those students whose "formula score" is considerably lower than their "number right score," and who fail as many difficult items as easy ones.

The task, in this test, is to start a hypothetical trip from a specified point of departure within an outlined square of a grid, and, after following simple successive directions, to indicate the point of arrival by reference to the perpendicular axis of the grid. A sample item follows.

DIRECTION
FINDER 1.

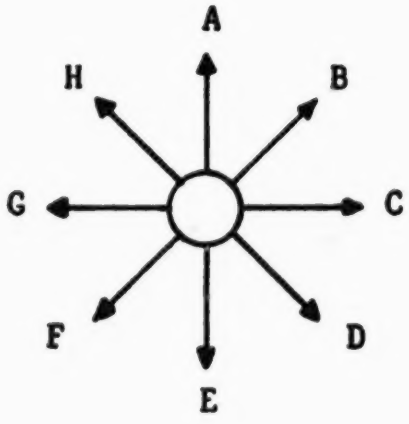
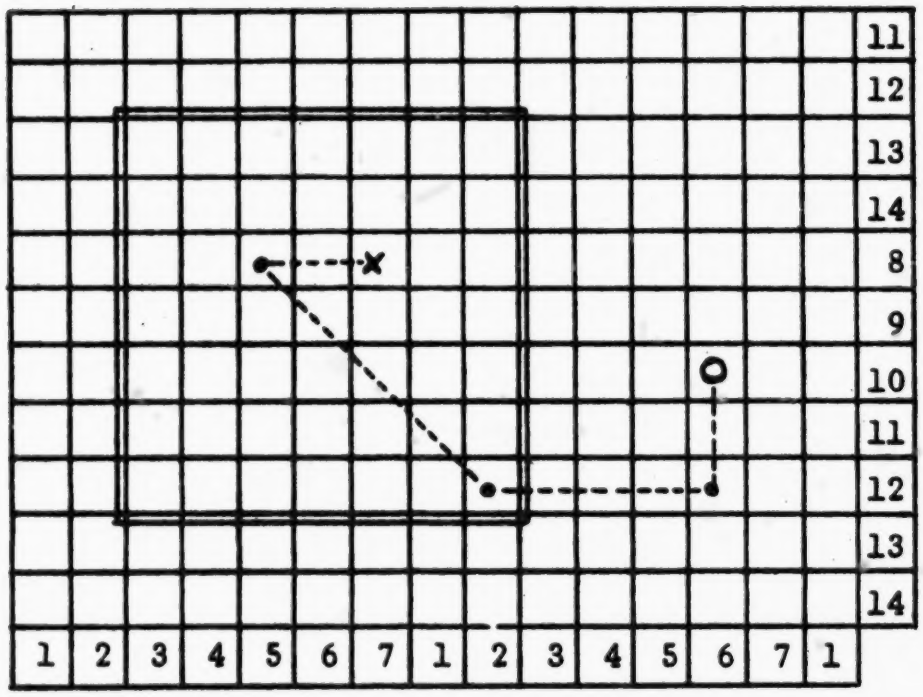


CHART (Sample)



Start at 7-8
Use Direction Finder 1

Move 2 spaces toward G
Move 4 spaces toward D
Move 4 spaces toward C
Move 2 spaces toward A

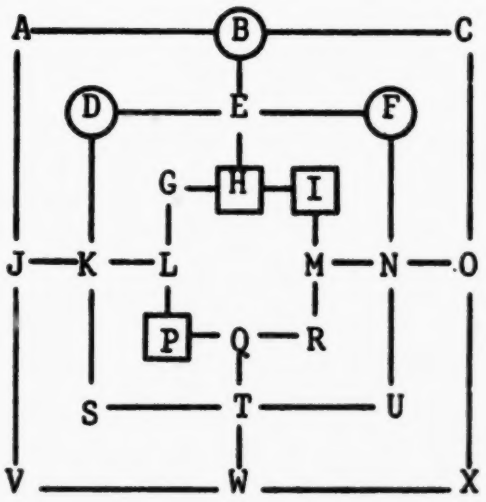
Answer: 6-10.

The test—a semi-verbal one—has a total of twelve items (maximum score: 24), and it takes thirty-two minutes to administer.

Game Playing Test

This also is a new test in the field of mental measurement. It was designed to be a measure of the degree of foresight and planning one employs in arriving at a specified objective. This test is expected to tap the factors underlying the concept of "planning ahead." In the field of automotive mechanics, as well as in other technical fields, individuals who take time to plan ahead, and take precautions for any foreseeable problems, are considered by service managers to be better mechanics than those who do not. Their respective pay-checks reflected these traits (from Job Analyses). If the test does indeed turn out to measure some of these factors, it will be considered as a significant contribution in the field of psychological measurement.

This test is also nonverbal, except for the somewhat lengthy directions. The subject is asked to select from among four alternatives the move which is critical in allowing him to win a game in spite of the greater versatility of moves available to his opponent. The chips are arranged on the grid in such a way that the subject—having the squared chips—always has a chance to win if he makes the correct move. A sample item follows.



To win this game:

A Which straight line must you make?	B Which is your first move?	C How many moves are required?
(a) G-L-P	(a) from H to E+	(a) two
(b) I-M-R	(b) from P to L	(b) four+
(c) P-Q-R	(c) from I to M	(c) five
(d) G-H-I+	(d) from H to G	(d) six

The object of this game is to complete a straight line (three letters in a row) with the fewest moves and before the opponent does—following all the rules of the game: one move at a time; must move when the turn comes; chips move only along straight lines; no skipping of unoccupied (letter) stations, nor jumping over occupied ones. For each figure, the subject must indicate on his answer sheet the straight line, the first move, and the number of moves included in his completed strategy.

The test has eleven figures with a total of thirty-three items, and it takes thirty-five minutes to administer. (A figure is a game lay-out.)

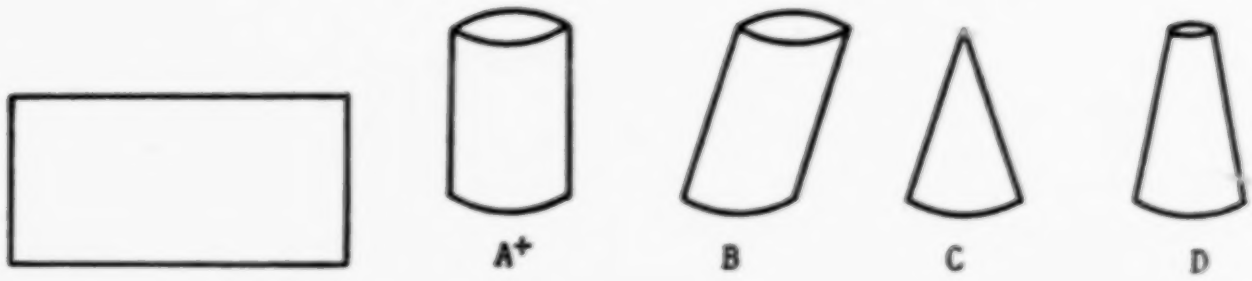
One of the unusual characteristics of this test is the motivational properties it possesses. Once they start, students become intensively involved in the task—an involvement reminiscent of programmed instruction.

Survey Of Object Visualization

This is a relatively unknown test published by the California Test Bureau.* It was added to the battery as a supplement for use in the area of spatial relations which is involved in a wide variety of technical, mechanical, drafting, and engineering occupations. It was decided to incorporate this already developed test instead of developing our own along the same traditionally popular lines.

*Miller, D. R. Survey of Object Visualization, Monterey, California: California Test Bureau, 1955.

The test was designed to measure the ability to visualize the manipulation of figures in space. The subject is asked to select among four alternatives, the three-dimensional object that will be made out of a two-dimensional unfolded design. A sample item follows.



The test has a total of forty-four items and it takes twenty-eight minutes to administer.

Biographical Information Blank

This is an inventory of items of biographical nature concerning the individual and his family. Some of these deal with the experience the person has had, others with his past and present interests, and still others with hig educational-vocational plans. The individual checks those items that are characteristic of himself (64, p. 2).

The usefulness of biographical information for the prediction of vocational choices was briefly described in the previous chapter, in connection with the Connecticut Vocational Schools study and the job analyses (p. 5), and it is extensively discussed in the literature (94, 104).

The blank contains a large number of groups of related items, and it has no set time limit for its administration, but usually requires about one-half hour to complete.

B. TEST MANUAL

III. ADMINISTRATION AND SCORING

GENERAL STATEMENT

A number of administrators throughout the nation, the extent of whose previous testing experience encompassed a wide range from none to considerable, have administered the tests in the battery with a minimum of difficulty. This fact alone attests to the probability that competent professionals, counselors, and teachers, who become thoroughly acquainted with the details of its administration, can give and score the National Aptitude Survey. Preview of this section as well as intermittent rereadings, are necessary, however, for the attainment of this objective.

THE MATERIALS NEEDED FOR THE TEST

Test Booklets

The seven tests that constitute the National Aptitude Survey appear in six separate booklets. These booklets bear the titles:

Reasoning Test
Vocabulary Test: General (odd items) Automotive (even items)
Game Playing Test
Chart Plotting Test
Survey of Object Visualization (California Test Bureau)
Biographical Information Blank (BIB)

Five of the six booklets are reusable any number of times. The students are not to make marks in the booklets and should receive instructions to this effect. In a majority of cases, under conditions of proper administration, students will follow instructions. It is possible that in a few instances some testees, for a variety of reasons, may mark in a test booklet itself, rather than on the answer sheet. Therefore, it is recommended that, a short time after the completion of the test program, an official scrutinize each booklet for markings, and erase those detected. If such a precaution fails, it is best to replace the damaged booklet with a new one.

The sixth booklet, the BIB, is not reusable. Testees mark their answers in the space provided.

Answers Sheet

Administration of the complete battery requires two answer sheets, each of which can be scored by hand or by an IBM test scoring machine. The Survey

of Object Visualization uses an answer sheet available from the publisher, and the remaining set of tests requires an IBM answer sheet developed for the original administration of the battery. One side of this sheet has sufficient answer categories to satisfy the requirements of the Vocabulary and Reasoning tests; on its reverse side are the answer blanks for the other two tests.

Scoring Keys

Two scoring keys recommended for use with the test battery are the Number Right score and the Formula Score. There are hand-scoring stencils available which fit the IBM answer sheets. The Survey of Object Visualization requires a separate stencil.

Those answer sheets to be scored by IBM machine will require a machine-scoring stencil. The California Test Bureau will machine score answer sheets for the Survey of Object Visualization.

Suggestions For Scheduling Testing Programs

Administration of the test battery requires at most two hours and forty-five minutes. This time allotment includes a brief relaxation period after the second test, another after the fourth test, and another after the fifth. It also includes the brief period of time required to distribute the test booklets for each test.

GENERAL CONSIDERATIONS FOR ADMINISTRATION

Timing

The development of the test battery required the determination of the approximate time necessary to obtain meaningful scores on the test. Administrators of the test battery should adhere closely to these established time periods. At least one of the following devices is necessary: a stop watch or a wall clock with a sweep-second hand.

Preparing Materials

A common procedure is to distribute each booklet at the right of a class and have it passed to the left after use. Prior assemblage of the correct number of booklets and answer sheets should minimize difficulty in distribution. The first four tests together require but one answer sheet.

Other needed materials include extra copies of the test booklet, a supply of sharpened pencils (the class size plus fifteen percent), and extra answer sheets.

STEPS IN ADMINISTERING THE TESTS

Step 1. Introduction

The National Aptitude Survey is important for those ninth-grade students who are to enter the world of work immediately after graduation from high school. The examiner should introduce himself, then stress the personal value the test can have for most students in terms of kinds of jobs they might enjoy and be successful in. As the directions for each test are quite clear, there is little need to state more than a few simple words as to what the test is about.

Step 2. Pencils and Seating

Explain the procedure to follow: Should a student need a different pencil, he need merely raise his hand and the proctor will bring another one.

Ensure that the students have an adequate surface on which to write. A table is preferable to a tablet-arm chair; however the latter will serve adequately in most instances. The room arrangement should be such that there is ample space between the testees.

Step 3. Answer Sheets

After the distribution of the IBM answer sheet, the administrator should show the group, or dictate, the information that should be on the answer sheet. In some cases it will be necessary to explain the use of the IBM sheets. One might say,

> "In the tests which are to follow, the questions appear in special booklets. (Show a test booklet.) You are not to write in the booklet but are to mark your answers on the answer sheet that you have received. You are to indicate a choice of answer with a heavy black mark in the space that corresponds to your choice, like this. (Demonstrate at the board how to mark an answer on a sheet.) Make your marks black; fill the space between the two lines; erase any marks other than your answers."

Also announce: "Do not open these booklets until I tell you to do so." Check to see that after each distribution of booklets the students follow this instruction.

Step 4. Reading Directions

The test booklets contain the directions for each test. (Note: The directions for the Survey of Object Visualization appear on page 8 of the manual that accompanies that test.) The examiner is to read these aloud as the students read them silently. Additional instructions which do not appear in the test booklets are to be read aloud distinctly.

Step 5. Timing

Appropriate time limits for the tests appear below. The column headed Testing Time lists the recommended time for administration of each specific test. The minutes allocated for Instructions are suggested amounts and may vary with the circumstances. Excess expenditure of time on the Instructions may lengthen the period of time required for administration of the entire battery and cause scheduling conflicts.

TABLE 2

TIME LIMITS FOR ADMINISTRATION OF NATIONAL APTITUDE SURVEY

Tests	Time Limits		
	Instructions	Testing Time	Total
1. Reasoning Test	8 min.	15 min.	23 min.
2. Vocabulary Test	3 min.	20 min.	23 min.
3. Game Playing Test	15 min.	20 min.	35 min.
4. Chart Plotting Test	17 min.	15 min.	32 min.
5. Survey of Object Visualization (California Test Bureau)	3 min.	25 min.	28 min.
6. Biographical Information Blank		(no time limit)	

Step 6. Collecting Materials

After the Vocabulary, Reasoning, and Game-Playing tests, collect the booklets but not the answer sheet. At the conclusion of the Chart-Plotting test collect both the booklet and the answer sheet. Do the same after administration of the Survey of Object Visualization. Collect the Biographical Information Blank after 90% or more of the students have completed it.

The procedure employed during the development of the battery was to have each student pass his booklet to the left to be collected by the last individual in the row farthest to the left. Consistency in whatever method used is desirable.

It is advisable after the session for the person-in-charge to have a log prepared. This log could contain such information as size and identification of group tested, the room, date, proctors, deviations from regular procedures (e.g., only certain tests given), and names of those tested whose scores may require special attention.

IV. NORMS AND PROFILES

GENERAL STATEMENT

The National Aptitude Survey underwent rigorous standardization in two common meanings of the term: a prior phase, usually designated as the period of "experimental control," encompassed the development of the procedures outlined more fully in the previous chapters; during the phase that followed the administration of the test nationally, occurred the compilation of norms for the battery. Such norms can serve to locate a particular score by relating it to the scores of the thousands of other students who have also taken the same tests. Included in this booklet are frequency distributions for the raw scores (Tables 3 and 4), as well as percentile tables that constitute the norms for each test, based on samples of approximately equivalent size (Figures 1 and 2).

READING THE SCORE-DISTRIBUTION TABLES

The frequency distribution of the scores obtained by the actual participants on each of the tests appears in summary form in Tables 3 and 4. To determine how many participants obtained a certain score on a particular test, the reader should look at the table appropriate for the sample (Table 3 lists scores for the probability sample and Table 4 lists scores for the non-probability). At the bottom of the table is the key that gives the column number for any one of the tests. This key also provides the number of participants whose scores appear under that column heading. In the left- and right-hand marginal columns appear the raw scores. One finds in either column the raw score that interests him, then moves across the table to the entry under the proper column heading, e.g., 278 participants out of 6642 obtained a raw score of 15 on the Survey of Object Visualization (column 6).

READING THE NORM TABLES

The National Aptitude Survey provides norms expressed as percentile ranks to facilitate the interpretation of scores on the test. There are two sets of norms, each based on a different sample population. One set (Figure 1) derives from the larger of the two, the probability sample; the other (Figure 2) from the substantially smaller non-probability sample.

The percentile ranks appear in two modes. In the first (Figure 1), the percentile ranks are located in the first and last columns and the raw scores fall under the column headings which designate the relevant test; in the second instance (Figure 2), the range of raw scores borders the left-hand column of the sheet, and the percentile ranks fall within the corresponding test columns.

The use of either figure requires knowledge of an individual's score on a particular test. To use Figure 1, the interpreter locates the raw score in the column labelled with the name of the test on which the score was obtained. The percentile rank which borders the extreme columns on either side, on a horizontal line with the raw score, is the rank sought. This numerical figure states what percentage of the norm group attained a score equal to the raw score which

TABLE 3

FREQUENCY DISTRIBUTION OF RAW SCORES ON TESTS OF THE NATIONAL APTITUDE SURVEY (PROBABILITY SAMPLE)

Raw Score	1	2	3	4	5	6	Raw Score
44						0	44
43						4	43
42						1	42
41						11	41
40						19	40
39						21	39
38						12	38
37						29	37
36						45	36
35						58	35
34						13	34
33					4	73	33
32					9	60	32
31					15	106	31
30	40	269	3		26	29	30
29	93	440	22		39	110	29
28	157	461	38		63	143	28
27	233	440	57		94	149	27
26	207	305	94		123	39	26
25	310	300	106		150	178	25
24	344	371	141	964	205	152	24
23	361	354	154	695	195	193	23
22	361	325	173	602	266	28	22
21	365	277	219	460	310	187	21
20	385	257	256	407	356	228	20
19	366	263	304	326	411	197	19
18	409	216	335	352	425	42	18
17	379	233	353	273	459	236	17
16	354	202	386	272	429	230	16
15	365	193	425	191	450	278	15
14	362	160	396	206	442	45	14
13	307	167	474	167	408	273	13
12	290	155	495	166	367	269	12
11	255	164	416	109	325	253	11
10	186	166	411	133	274	40	10
9	153	143	343	110	222	250	9
8	120	145	318	125	183	273	8
7	88	148	246	110	153	316	7
6	43	111	185	138	111	52	6
5	28	75	120	139	62	281	5
4	20	72	104	162	31	293	4
3	3	52	54	134	21	277	3
2	2	27	24	125	10	48	2
1	1	11	11	83	11	248	1
0	1	7	4	46	5	631	0

Key to Table 3

Col. No.	Name of Test	N	Col. No.	Name of Test	N
1	Reasoning	6668	4	Chart Plotting	6617
2	Vocabulary: General	6669	5	Game Playing	6654
3	Vocabulary: Automotive	6669	6	Survey of Object Visualization	6642

TABLE 4

FREQUENCY DISTRIBUTION OF RAW SCORES ON TESTS OF THE NATIONAL APTITUDE SURVEY (NON-PROBABILITY SAMPLE)

Raw Score	1	2	3	4	5	6	Raw Score
44						0	44
43						0	43
42						0	42
41						0	41
40						2	40
39						6	39
38						3	38
37						4	37
36						11	36
35						7	35
34						1	34
33					0	8	33
32					1	17	32
31					0	14	31
30	4	11	6		0	1	30
29	9	43	15		4	14	29
28	8	40	20		7	23	28
27	25	45	46		11	21	27
26	26	49	45		14	5	26
25	40	51	55		8	26	25
24	42	52	42	124	21	23	24
23	35	40	42	86	25	24	23
22	48	41	48	98	36	7	22
21	41	39	41	65	26	21	21
20	53	35	38	70	32	38	20
19	44	30	35	40	48	26	19
18	39	23	41	42	51	9	18
17	44	37	40	29	47	38	17
16	45	20	27	33	47	22	16
15	50	18	26	19	55	18	15
14	40	21	29	12	52	6	14
13	26	21	26	12	54	23	13
12	26	19	21	11	31	25	12
11	27	11	21	10	33	30	11
10	17	18	18	7	34	4	10
9	17	12	14	7	29	20	9
8	11	17	13	3	24	30	8
7	8	12	8	7	12	19	7
6	8	10	10	10	11	4	6
5	6	10	4	10	8	27	5
4	1	7	1	11	3	21	4
3	1	3	1	9	1	21	3
2	1	4	1	10	3	4	2
1	0	3	0	3	1	23	1
0	0	1	0	2	0	65	0

Key to Table 4

Col. No.	Name of Test	N	Col. No.	Name of Test	N
1	Reasoning	744	4	Chart Plotting	732
2	Vocabulary: General	743	5	Game Playing	731
3	Vocabulary: Automotive	742	6	Survey of Object Visualization	713

FIGURE 1

NATIONAL APTITUDE SURVEY

The University of Michigan - Automobile Manufacturers Association
School of Education, Ann Arbor, Michigan

NORMS
for
NATIONAL PROBABILITY
SAMPLE

NINTH GRADE BOYS

N: 6694 ±

RAW SCORES

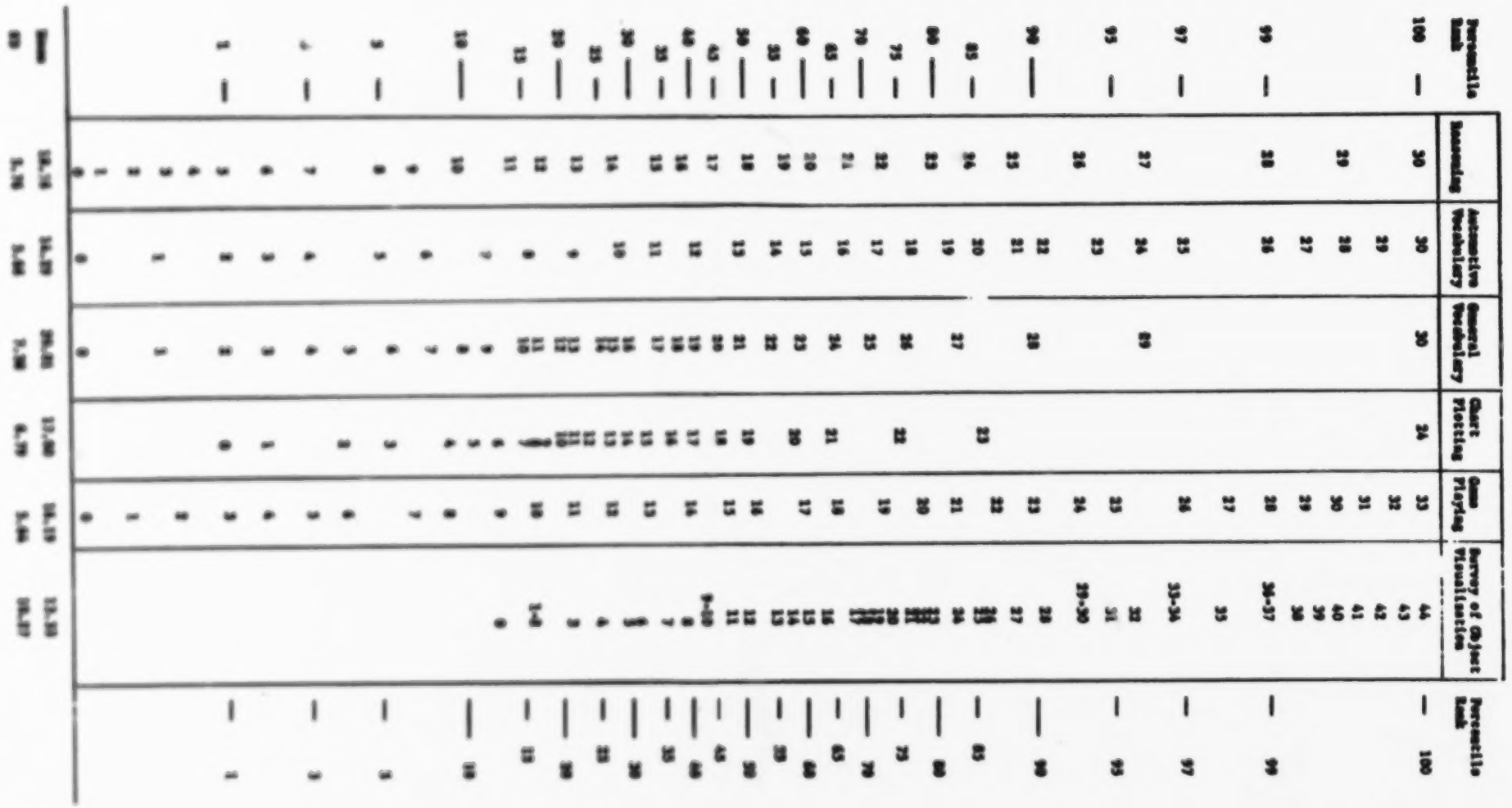


FIGURE 2

NATIONAL APTITUDE SURVEY

The University of Michigan - Automobile Manufacturers Association
School of Education, Ann Arbor, Michigan

TENTH GRADE BOYS

N: 744

NORMS
for
NON-PROBABILITY
SAMPLE

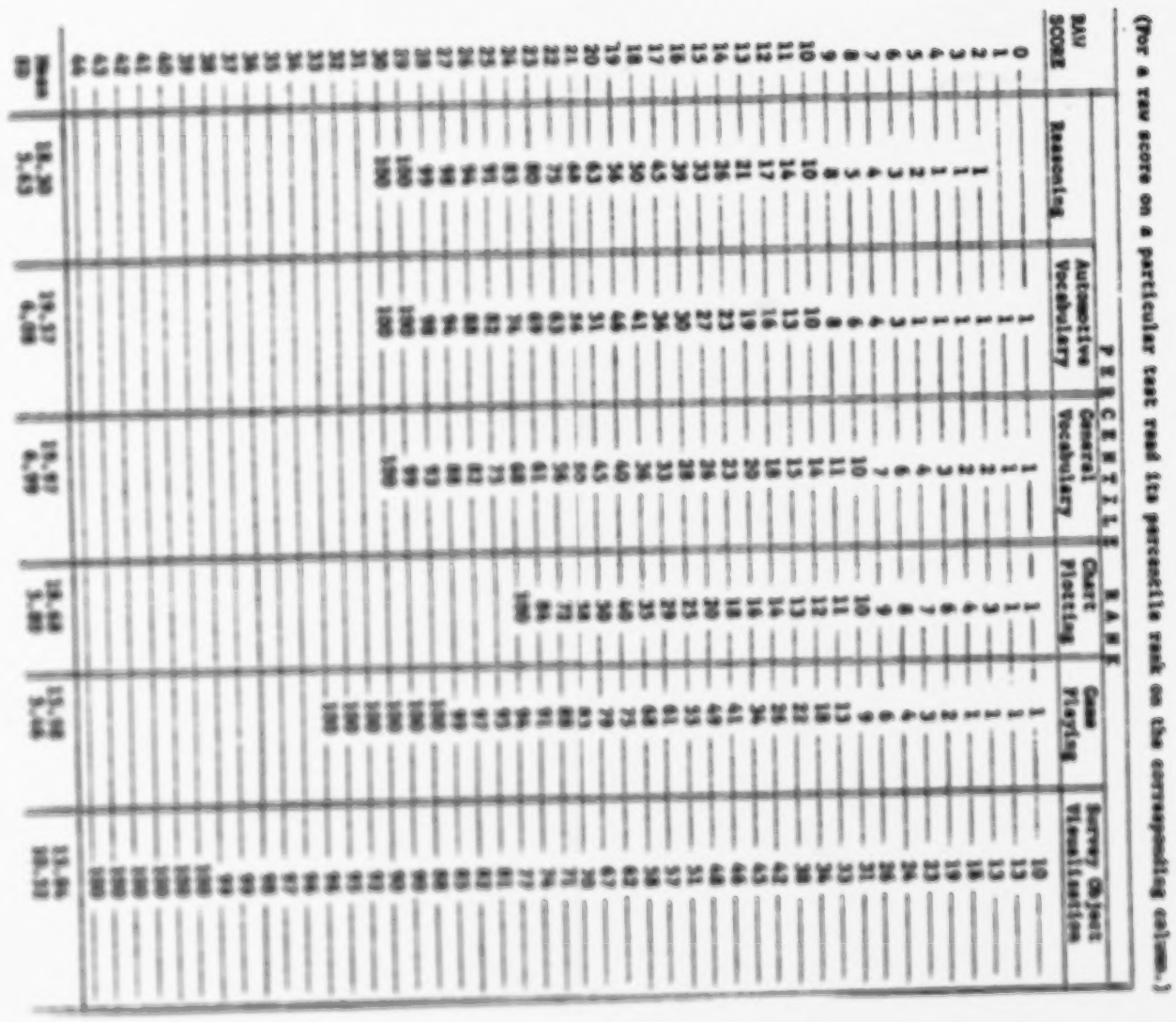


FIGURE 3

INDIVIDUAL PROFILE FORM

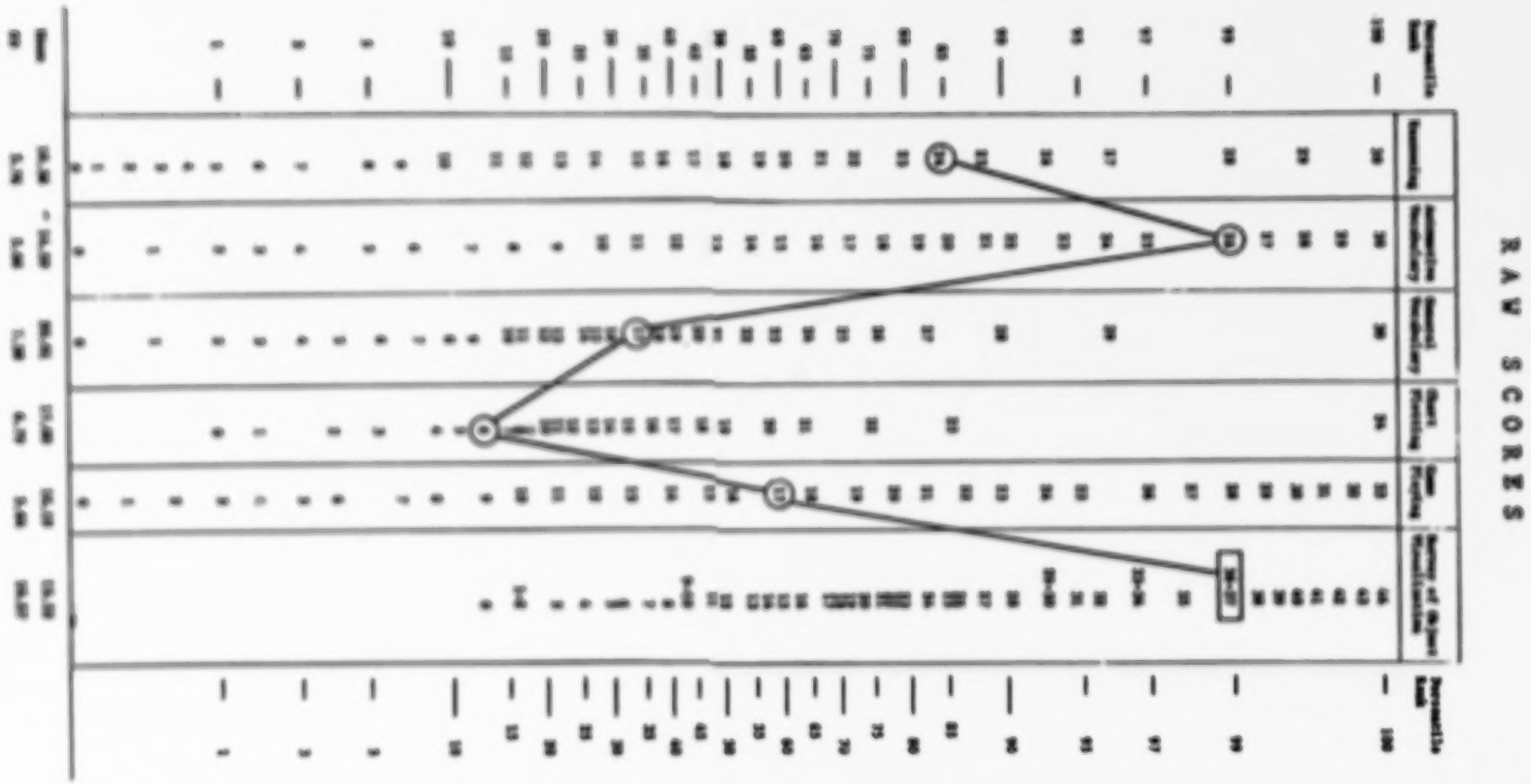
NATIONAL APTITUDE SURVEY

The University of Michigan - Automobile Manufacturers Association
School of Education, Ann Arbor, Michigan

SAMPLE

Name: JAMES TAYLOR Age of Boy: 14 Years 6 Months

Place of Testing: GREENFIELD HIGH SCHOOL Norm: GRADE 9 - BOYS Date of Testing: 10-15-1961



is of concern. To use Figure 2, the procedure is reversed. One seeks the raw score in the leftmost column, and, on a horizontal line with it under the pertinent test heading, locates the percentile rank in question, and repeats this procedure as many times as there are scores. The final step, in the case of more than one score for a person, is to encircle the raw scores or percentile ranks, and then to connect them by means of straight lines. Figure 3 represents a hypothetical test profile based on several scores.

During the interpretation of the profile, the question may arise as to how large a difference between two scores is statistically significant. The placement of raw scores or percentile ranks in the figures permits no direct comparison for this purpose.* A form of metric conversion, although helpful, is not available at this time. In general, high scores on one test are superior to low scores on others; for those scores whose differences are not extreme, other background information should contribute to the judgment.

The order of the columns themselves has no significance. Their positions relative to one another imply no rank order of importance in either direction. The juxtaposition of any two columns is a matter of administrative convenience alone; any other order would have served the purpose equally well.

INTERCORRELATIONS OF THE TESTS

A usual practice in the construction of a test battery is to include only those tests which have little correlation one with another. The rationale for it is this: If one test correlates highly with another in the same battery, then it can substitute for the latter. Mutually interchangeable tests duplicate but contribute little else to the variance to be explained. The National Aptitude Survey tests avoid duplication wherever it is possible, and its component tests intercorrelate from .45 to .14. Ideally those tests that constitute the final battery should correlate with each other not at all; in practice the ideal is rare, and the subjective decision as to what is acceptable serves to satisfy this requirement.

In Table 5 appear the intercorrelations of all the tests. Kaltsounis has provided the data in his doctoral dissertation. He selected (from the entire probability sample) a sample of (N = 1032) individuals in order to make a concentrated study of the relationship between the student's scores on the Survey tests and their performance in high school vocational education courses. One of his twelve validational sub-groups had taken for credit a course entitled Auto Shop. Table 5 presents as well for comparison the intercorrelations of the tests for this smaller group.

---

*The Differential Aptitude Test Battery has a convenient method for locating significant differences. The vertical differences between scores that measure more than one inch can with assurance be considered statistically significant.

TABLE 5

INTERCORRELATIONS OF ALL TESTS OF THE NATIONAL APTITUDE SURVEY
FOR THE ENTIRE STUDY SAMPLE (N = 1032)
AND THE AUTO SHOP GROUP (N = 31) (KALTSOUNIS' STUDY).

	Test	Sample	1	2	3	4	5	6
1.	Reasoning	St. Sample	-	.33	.14	.32	.33	.43
		Auto Shop	-	.18	.18	.29	.35	.48
2.	Gen. Vocab.	St. Sample		-	.38	.39	.35	.31
		Auto Shop		-	.43	.30	.41	-.11
3.	Auto. Vocab.	St. Sample			-	.22	.15	.28
		Auto Shop			-	.19	.38	.03
4.	Chart Plot.	St. Sample				-	.45	.39
		Auto Shop				-	.37	.20
5.	Game Playing	St. Sample					-	.40
		Auto Shop					-	.35
6.	S. Obj. Vis.	St. Sample						-
		Auto Shop						-
N 1032	Mean	St. Sample	18.75	20.36	14.65	17.25	16.24	20.85
N 31		Auto Shop	15.74	18.13	13.94	15.03	13.71	18.61
N 1032	S.D.	St. Sample	5.62	7.17	5.67	6.73	5.51	7.68
N 31		Auto Shop	5.05	7.15	6.91	7.56	4.14	6.68

V. RELIABILITY

GENERAL STATEMENT

The constancy of a test over time is one of its most important characteristics. This consistent aspect, known as its reliability, is the extent to which a set of measurements is free from random-error variance. It is meaningful in that it provides an estimate of the importance of its systematic variance in relation to its entire variance (51).

The problem that faces most test builders is to maintain a balance between the reliability of a test and the limits for its administration. The National Aptitude Survey features satisfactory coefficients of reliability for nearly all of its tests, yet meets the requirements of length necessitated by the school's schedule of classes.

TYPES OF RELIABILITY COEFFICIENTS USED

The tests of the National Aptitude Survey are of the power type, as at least ninety percent of the participants were able to complete them. One estimate of reliability that appears to meet the requirements of a power test is the Kuder-Richardson Formula 21. Each test has its own reliability correlation coefficient (Table 6).

TABLE 6

ESTIMATED RELIABILITIES AND STANDARD ERRORS OF MEASUREMENT FOR THE NATIONAL APTITUDE SURVEY (PROBABILITY SAMPLE)

Test	r	S.E.M.
Reasoning Test	.81	2.5
General Vocabulary Test	.91	2.3
Automotive Vocabulary Test	.79	2.5
Game Playing Test	.76	2.8
Chart Plotting Test	.93	1.8
Survey of Object Visualization: Raw Score	a	b

a. Expected to be approximately .83 b. Expected to be approximately 4.1

For comparative purposes, a further analysis of the reliability of the tests appear in Table 7. These data came from Kaltsounis' dissertational study.

TABLE 7

COEFFICIENTS OF RELIABILITY AND OTHER STATISTICS ON THE NATIONAL APTITUDE SURVEY FOR THE STUDY SAMPLE AND ONE COEFFICIENT FOR THE NATIONAL SAMPLE

National Aptitude Survey	STUDY SAMPLE N	M	SD	SE of Meas.	Split-half[a]	Rulon	Kuder-Richardson No. 21	NATIONAL SAMPLE Kuder-Richardson No. 21
Reasoning	1032	18.57	5.62	2.51	.72	.71	.80	.81
General Vocabulary	1032	20.36	7.17	2.27	.92	.92	.90	.91
Automotive Vocabulary	1032	14.65	5.67	2.60	.80	.80	.79	.79
Chart Plotting	1032	17.25	6.73	1.78	.92	.92	.93	.93
Game Playing	1032	16.24	5.51	2.76	.65	.65	.75	.76
Survey of Obj. Visualization	1032	20.85	7.68	3.16	.86	.86	.83	—[b]

[a]Corrected by the Spearman-Brown formula.

[b]Not available; expected to be approximately .83.

FORMULAS USED:

(a) Estimate of Reliability (r): Kuder-Richardson, Formula 21 (Guilford, J.P., Fundamental Statistics in Psychology and Education. New York, McGraw Hill, 1956. 3rd edition, p.455 (17.19).)

$$r_{tt} = \frac{n\sigma^2_t - M(n - M)}{(n - 1)\,\sigma^2_t}$$

In which

r_{tt} = reliability estimate
n = number of items in test
σ^2_t = variance of the test
M = mean

(b) Standard Error of Measurement (S.E.M.): Ibid., p.441 (17.8).

$$\sigma_{tm} = \sigma_t \sqrt{1 - r_{tt}}$$

In which

r_{tt} = reliability estimate
σ_{tm} = standard error of measurement
σ_t = standard deviation of the obtained scores

An alternate method of interpreting a test's reliability is to make use of the standard error of measurement. This statistic indicates a pair of limits, one on either side of the obtained score. It is possible, then, to deduce with a specified degree of confidence, an interval that is sure to include the true score, or the obtained score less any random errors. Table 6 lists the standard errors of measurement that locate the points of the bracket around an obtained score.

VI. VALIDITY

GENERAL STATEMENT

The National Aptitude Survey research staff has concentrated on obtaining three kinds of validity estimates for the tests: face, concurrent, and predictive. The construction of the tests relied on job descriptions and previous research. Kaltsounis in his dissertation (63) predicted course grades for a group of students (N = 31) enrolled in Auto Shop from knowledge of their Survey scores. This chapter is an account of the procedures followed to determine the long-term predictive validity of the tests. It answers the question: Was it possible to discriminate potential automotive repairmen for a three-year period immediately following high school graduation from among other male students at the ninth or tenth grade level by means of the test scores?

VALIDATIONAL SAMPLE

The total sample of individuals numbers 7625. An appropriate representative sample of non-automotive repairmen (NAR) was obtained from this total to serve as a cross-validational group for the final analysis, and to determine which tests and/or inventory items should be used to specify the predictive validity of the Survey. A sample one-seventh the size of the total sample was considered appropriate to meet these needs.

The selection of the NAR representative sample (SUBSAM) proceeded in the following manner:

(a) Since each participant had received an identification number which indicated his school and number within the school, his number could occupy a rank in an ordinal sequence;

(b) one could select by means of a table of random numbers the starting number among the first seven; and

(c) those participants that occupied a position in the ordinal sequence that was a multiple or seven away from the first man would appear in the sample.

This method produced a sample of 1084 participants for the NAR representative sample.

The sample of automotive repairmen (MECSAM) consisted of those participants who had replied, in answer to the questionnaire for the third follow-up,

that they were working, or had worked, as mechanics.* As a check on their statements, they also provided a description of the work they performed. The clerical staff checked these descriptions and eliminated all mechanics not in the automotive repair field as described in the bibliographical reference 3. Although the third follow-up had contacted only eighty percent (80%) of the participants, it was still possible to refer to the second follow-up for information about an additional seventeen percent (17%) (17 of the 20% not contacted on the third follow-up), thus providing information on ninety-seven percent (97%) of the participants.

A comparison of the sample lists for MECSAM and SUBSAM disclosed that the random method of selection had resulted in the inclusion of forty-eight (48) automotive repairmen in SUBSAM. These forty-eight men were replaced with other non-automotive repairmen from the remaining population. One researcher chose at random a name from the middle of the list and selected every seventh name to identify the forty-eight substitutions. The procedure continued until SUBSAM contained no known automotive repairmen, and MECSAM all known repairmen.

A tabulation of scores for each test revealed that some (N = 84) lacked partial scores. None of the scores of these participants appeared in the analyses that followed. Table 8 summarized the adjustments in size of the respective samples.

TABLE 8

DISTRIBUTION OF PARTICIPANTS INTO RELEVANT SUBSAMPLES

Category		Auto Repairmen	Non-Automotive Repairmen (NAR)		Total
			Representative Sample (SUBSAM)	Remaining Population (NAR)-(SUBSAM)	
Total Sample Population N	(a)	387	1084	6154	7625
Incomplete Score N	(b)	11	25	148	184
Adjusted N	(c)	376	1059	6006	7441

*See page 9, Longitudinal Aspects of the Sample.

To test the assumption that SUBSAM (N = 1059) was representative of the entire population of NAR (N = 7065), Perkuchin applied a stepwise linear regression analysis* to learn whether the test scores would in fact discriminate the SUBSAM group from the remaining NAR group. The assumption of sample representativeness appeared tenable: the correlation between a man's group membership and his best score was .02 and the multiple linear correlation coefficient between group membership and the six scores was .04.

ANALYSIS OF TEST SCORES OF THE VALIDATIONAL SAMPLES

For each participant there were four hundred six (406) measured variables. The greater number (N=400) of the variables consisted of inventory items. A comparison, MECSAM vs. SUBSAM, of the frequency distribution of responses for each item showed that fifty-two (52) discriminated to some degree between the groups.

The refined list of variables now included the six (6) tests and fifty-two (52) inventory items, for a total of fifty-eight (58). A second regression analysis, applied to the fifty-eight (58) variables, designated two test scores and six inventory items as contributing most in maximizing the discrimination between members of MECSAM and SUBSAM. The size of the samples (MECSAM and SUBSAM) used in this analysis was six-sevenths of the stated size; the remaining seventh served as the cross-validational group. Selection of the fractional parts of the two samples was by the method outlined earlier (p. 35). The actual size of the groups were respectively 322 and 907 for the larger fraction, and 54 and 152 for the smaller. Table 9 lists the order in which the eight variables appeared in the regression analysis. It also furnishes the multiple linear correlation coefficients and the coefficients of determination for each of the eight steps.

Only eight variables appear in this table because the contribution of the ninth variable, as well as the forty-three (43) that remained, was insignificant relative to the contribution of the first eight. The last coefficient of multiple determination (hereafter referred to as the coefficient of determination) in that table, i.e., .197, indicates the proportion of variance in the respective participants' occupations that is dependent upon, associated with, or predicted by the eight variables combined with the regression weights used. For this regression analysis a participant's occupation falls into one of two categories: he is (MECSAM) or he is not (SUBSAM) an automotive repairman. Whatever it is that the two test scores and the six inventory items in weighted combination measure, in this analysis it accounts for 19.7% of the variance in the respective participants' occupations. They do not account for the remaining 80.3% of the variance.

---

*Linear Stepwise Regression Analysis Program, The University of Michigan Computing Center, 1967.

TABLE 9

MULTIPLE CORRELATION COEFFICIENTS AND COEFFICIENTS OF DETERMINATION FOR EIGHT SELECTED VARIABLES (322 MECHAM VS.907 SUBSAM)

Regression Step No.*	Variable	Multiple Correlation Coefficient	Coefficient of Multiple Determination
	Survey Test		
1.	Automotive Vocabulary:	.328	.108
2.	General Vocabulary:	.380	.145
	Biographical Inventory		
3.	Item A**	.396	.157
4.	Item B	.410	.168
5.	Item C	.420	.176
6.	Item D	.428	.184
7.	Item E	.437	.191
8.	Item F	.444	.197

*The variable listed after each Step No. is the variable combined with the variables in the steps preceding the one considered. All of these variables produce the listed multiple correlation coefficient and coefficient of multiple determination for the respective Step No.

**The items in the actual inventory were unnumbered consecutively. Appendix B contains the complete item and its alternatives.

In the determination of the representativeness of SUBSAM, a comparison with the remaining NAR in regard to inventory items was left pending, since one reason for taking a subsample was to defer coding several thousand lengthy inventories. The regression analysis indicated that only six items had value. Also no one inventory item accounted for more than 10.8% of the variance in participants' occupations and no two accounted for more than 14.3% of the variance. If the inventory items had accounted for more variance, then they would have replaced the two tests in the first two steps of the regression analysis. The researchers sought to determine if their value was such as to be worth the additional coding. The decision was to use a graphic visual comparison in order to perceive what their inclusion was worth.

The decision whether to retain or eliminate the six items depended on a consideration of their effect on prediction. If their inclusion did not produce an estimate of the chances a boy has of becoming an automotive repairman significantly different from an estimate made with them excluded, there would be little reason to perform the extra calculations, for the two test scores alone would be as useful.

Perkuchin used percentile rank order information* to compare both estimates; one estimate was derived from all eight variables and the other derived from the two tests alone. He proceeded thus:

(a) he computed a composite score** for MECSAM (N=322) and for SUBSAM (N=907);
(b) the obtained composite score was rank ordered with the highest numerical value placed in the first position;
(c) the rank order facilitated the division of scores into a percentile rank order with fourteen percentile ranges of various sizes selected (see col. 3, Table 10);
(d) he then computed the proportion of automotive repairmen within each percentile range;
(e) he repeated the procedure for the cross-validational groups, MECSAM (N=54) and SUBSAM (N=152).

For both conditions the proportions from the cross-validational sample were similar to those obtained with the larger groups.

---

*Percentile Range Analysis Program, Programmer Dan N. Perkuchin, July 1967.

**The composite score is the predicted score computed by the regression program for the respective number of variables, eight or two. The program computes a constant term and a coefficient for each variable to maximize the discrimination between both groups. To obtain a participant's composite score it is necessary to multiply the test scores and the responses to the inventory items by the proper coefficients (a, b, . . . n) and to add the results algebraically to the constant term:

composite score = constant + a(automobile vocabulary score) +b (general vocabulary score) +...+n(last inventory item)

TABLE 10

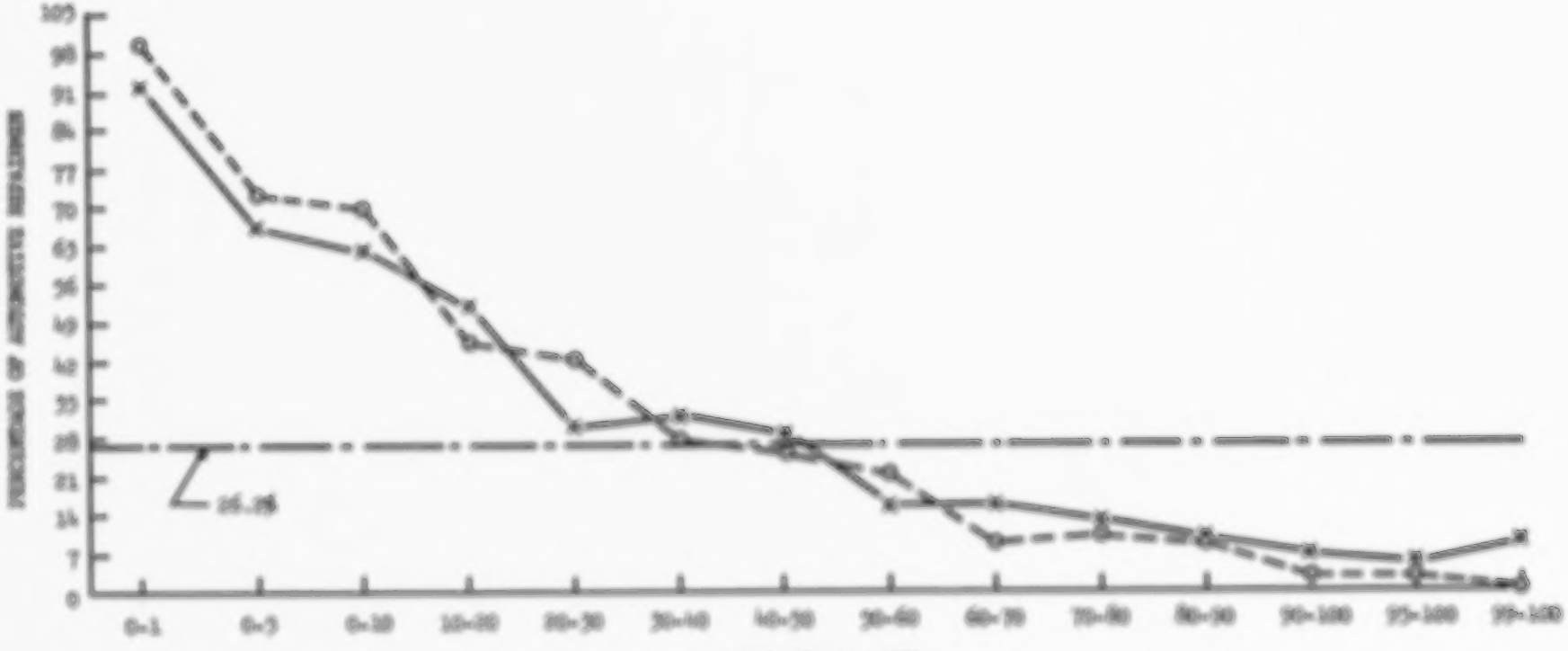
PERCENTILE RANGE AND COMPOSITE SCORE BOUNDARIES, PROPORTIONS OF AUTOMOTIVE REPAIRMEN, AND LIKELIHOOD ESTIMATES OF BECOMING AN AUTOMOTIVE REPAIRMAN

Composite Score Range	Percentile Range	Proportions of Automotive Repairmen*	Chances of Becoming an Automotive Repairman**
Highest Score - Lowest Score	0 to 100	322/6328 = .051	5/100
1.8465 - less than	99 to 100	16/63 = .254	25/100
1.8707 - less than	95 to 100	64/315 = .203	20/100
1.8855 - less than	90 to 100	108/633 = .171	17/100
1.9097 - 1.8856	80 to 90	56/632 = .089	9/100
1.9276 - 1.9098	70 to 80	42/639 = .066	7/100
1.9413 - 1.9277	60 to 70	36/622 = .058	6/100
1.9529 - 1.9414	50 to 60	17/638 = .027	3/100
1.9645 - 1.9530	40 to 50	17/640 = .027	3/100
1.9761 - 1.9646	30 to 40	19/626 = .030	3/100
1.9877 - 1.9762	20 to 30	10/622 = .016	2/100
2.0025 - 1.9878	10 to 20	9/644 = .014	1/100
greater than - 2.0026	0 to 10	8/632 = .013	1/100
greater than - 2.0141	0 to 5	3/316 = .009	1/100
greater than - 2.0351	0 to 1	1/64 = .016	2/100

*The proportion of automotive repairmen within a percentile range is also the probability of a boy, whose composite score falls within that range, becoming an automotive repairman within the three-year period after his expected class graduation date.

**This column is a restatement of the adjacent column after rounding the decimals to two places. Seventeen out of one hundred boys who have a composite score less than or equal to 1.8855 are likely to be automotive repairmen within a three-year period after their expected class graduation date.

Figure 4 shows the lines joining the percentages (proportions) of automotive repairmen within each percentile range computed with eight and two variables. Visual inspection reveals that the two lines do not differ appreciably. Because of this, the research team elected to eliminate the six items from the instruments needed to predict entrance into the field of automotive repair.



		Group ③												
X ①	Ratio	202/1209	11/12	42/63	75/120	64/124	77/123	39/120	35/123	19/124	19/123	15/124	11/121	8/123	3/61	1/13
	%	16.2	91.7	66.7	61.5	51.6	30.1	32.0	28.5	15.3	15.4	12.1	9.1	6.5	4.9	7.7
O ②	%	16.2	100.0	72.1	69.9	45.5	48.3	27.6	25.2	11.3	8.9	9.0	8.9	2.4	1.6	0.0
	Ratio	202/1209	12/12	44/61	86/123	56/123	59/123	34/123	31/123	14/124	11/123	12/123	11/123	3/123	1/61	0/12

① X represents the percentages computed from only two variables; i.e., Automotive and General Vocabulary tests.

② O represents the percentages computed from the eight variables; i.e., Automotive and General Vocabulary tests plus six inventory items.

③ Both conditions with the same participants (N = 1209; 202 automotive repairmen plus 907 non-automotive repairmen).

Figure 4. A comparison of percentages of automotive repairmen within percentile ranges computed from eight variables and from two variables.

Having decided that the two test variables contributed the most in maximizing the discrimination between members of MECSAM and SUBSAM, another regression analysis was used to obtain the multiple linear correlation coefficient, coefficient of determination, and the regression weights on these two variables for all the participants in the study. The appropriate groups for the analysis were the MECSAM automotive repairmen (N = 322) and the remaining NAR population (N = 6006); the remaining MECSAM members (N = 54) and SUBSAM (N = 1059) being the cross-validational groups.

In this regression analysis, the automotive vocabulary test and the general vocabulary combined, with their respective regression weights used, accounted for only 4.0% of the variance in the study's participants' occupations (322 MECSAM vs remaining NAR). The multiple correlation coefficient between the participants' occupations and the two test scores was .200.

Because 96.0% of the variance in participants' occupations is not accounted for by the variance that exists in the weighted combination of two test scores, the question arises as to what possible advantage one gains by giving the tests in the first place. The advantage lies in the information provided by the test scores. Within the continental United States approximately five percent (387/7625) of the male population, within three years after the date of expected high school graduation in this study, is now working at, or has worked in, some phase of automotive repair. With no knowledge whatsoever about a boy in the ninth or tenth grade in school, one can predict that his chances of entering the occupation of automotive repairman within three years after high school is one out of 20 (5/100). By means of a composite score computed from the automotive and general vocabulary tests the determination of a boy's chances of becoming an automotive repairman is improved over an estimate from no knowledge about the boy, even though these scores account for only 4.0% of the variance in the participants' occupations, the probabilities varying with the computed score. The likelihood estimate of a male ninth or tenth grade student becoming an automotive repairman within three years of graduation varies in accordance with the percentile range in which his score falls. Table 10 is a guide for locating the appropriate likelihood estimate for a score.

The description of the procedure for obtaining the percentage of automotive repairmen within a percentile range appears earlier in this chapter (p. 39). The determination of the entries—composite score boundaries, percentile range, etc.—for use in Table 10 followed a similar procedure: use of the stepwise linear regression program to maximize the difference between two groups; division of the samples to be compared into validative (6/7) and cross-validational (1/7) groups; and selection of sample members as before. The groups to be discriminated were six-sevenths of the automotive repairmen and the remaining NAR population. The regression equation provided by the program served to calculate the composite scores. The numerical value for the constant term

$$\text{Composite Score} = 2.0109259 - 0.00843046 \begin{matrix}\text{Automotive}\\ \text{Vocabulary}\\ \text{Test Score}\end{matrix} + 0.00315935 \begin{matrix}\text{General Vocabu-}\\ \text{lary Test Score}\end{matrix}$$

and the two coefficients used appear in the equation. Perbuchin's computer program rank-ordered the composite scores for the 6328 cases (i.e., MECSAM, N = 322, and the remaining NAR population, N = 6006) with the lowest scores ranked highest. From the computed percentile-rank order, composite-score boundaries for each percentile range were established and the proportion of automotive repairmen within each percentile range of all the men within that range was determined.

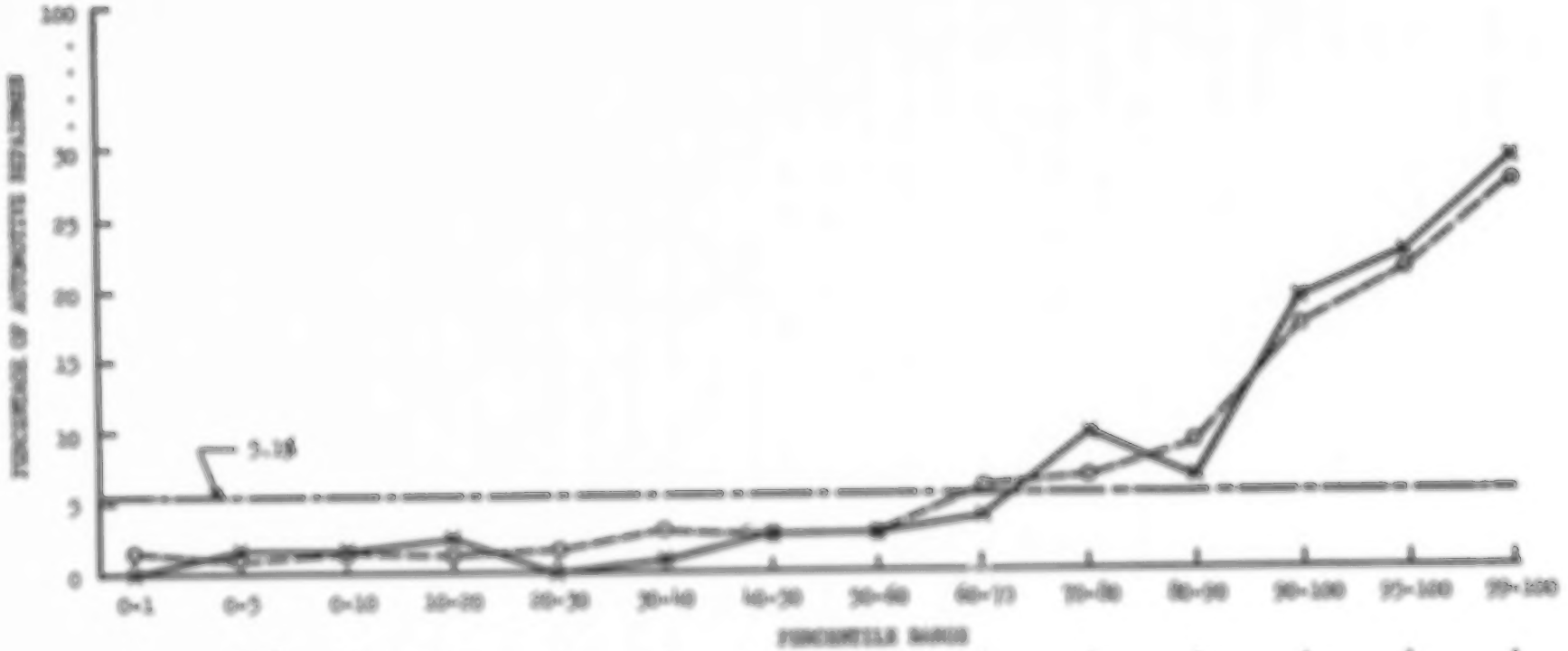
A visual comparison of the points that represent the percentage (proportion) of automotive repairmen in each range is possible by inspection of Figure 5. The points indicate the percentage of automotive repairmen obtained from the two sets of composite scores, one set obtained from the validational sample and the other from the cross-validational. Their proximity to one another emphasized by the curves would appear to sustain the belief that the proportion of automotive mechanics obtained from the larger group is representative of the entire ninth and tenth grade male population within the continental United States during 1960 and 1961.

SUMMARY

The validation population consisted of three groups: a group of automotive repairmen (N = 376), a representative sample of non-automotive repairmen (N = 1059), and the remaining group of non-automotive repairmen (N = 6006.) Recombined, they constitute the original representative sample of the ninth and tenth grade male population in the continental United States during 1960 and 1961. A comparison by linear regression of the six test scores of the two large samples revealed no significant differences between them on that factor. Therefore, in the subsequent analyses, the representative sample of non-automotive repairmen served as a substitute for the entire population of non-automotive repairmen.

In a comparison of the automotive repairmen with the representative sample, the regression analysis disclosed that the best predictors among those variables measured were the automotive and general vocabulary tests, and six biographical inventory items. Together they accounted for about twenty percent of the variance in participants' occupations. A comparison in predictions made by the above eight variables and the two test variables only resulted in the elimination of the inventory items. The two tests provided the scores needed to determine the percentage of automobile repairmen within various percentile ranges. This normative distribution lists the probability that a student with a particular score has of becoming an automotive repairman.

An additional product of the regression analysis was an equation to use in the computation of a score. A cross-validation of the equation on two subgroups served as support for its validity.



		Group													
0 ①	Ratio	502/6508	1/64	3/316	8/632	9/644	10/648	19/636	21/640	27/638	36/628	42/639	56/632	108/633	64/385	16/63
	%	7.1	1.6	0.9	1.3	1.4	1.6	3.0	3.7	3.7	5.8	6.6	8.9	17.1	20.3	25.4
X ②	%	4.9	0.0	1.4	1.6	2.3	0.0	0.9	2.7	2.7	3.8	3.9	6.6	15.8	22.6	28.6
	Ratio	54/1113	0/18	1/72	2/127	3/130	0/104	1/108	3/112	3/112	4/104	10/101	8/121	20/104	12/53	2/7

① 0 represents the validational group (N = 6508; 502 automotive repairmen plus 6006 non-automotive repairmen).

② X represents the cross validation group (N = 1113; 54 automotive repairmen plus 1059 non-automotive repairmen).

Figure 5. A comparison of percentages of automotive repairmen within percentile ranges for the validational group and a cross-validational group.

VII. USE OF THE VALIDATED TESTS

The staff recommends that the National Aptitude Survey in modified form be made available to public high schools. This revised battery might use the same format for the two best predictors as before: "a single booklet, under one set of directions, with the items alternating from general (odd) to automotive (even) vocabulary."* It might well include the six inventory items (Appendix B) as an aid in developing a local interest inventory. Should the guidance counselor prefer to use an inventory which does not make use of the six items, he should feel free to do so, since they contribute a negligible amount to prediction.

The steps required of a potential user of this guide to obtain a probability estimate of a boy being an automotive repairman three years after high school graduation follows:

1. The Automotive and General Vocabulary tests are intended to be administered to boys in their ninth or tenth grade in school.

2. The obtained scores for the boys are then placed into the equation for computing the composite score.

3. The composite score is then used to enter the proper row of Table 10.

4. Once the proper entry point has been determined the percentile range in which the boy's composite score falls and his probability of being an automotive repairman three years after graduation from high school can be found by looking in the appropriate column of Table 10.

In illustration, let's assume a boy has an Automotive Vocabulary test score of 15 and a General Vocabulary test score of 20. The composite score equation would be: COMPOSITE SCORE = 2.0109259 - (0.00843046)(15) + (0.00515935)(20). COMPOSITE SCORE = 1.94765600 or 1.9477.

This composite score is greater than 1.9413 but less than 1.9529. Thus, 1.9477 is contained in the percentile range 50-60. The boys probability of being an automotive repairman three years after high school graduation is .027.

The results shown in Table 10 indicate that the boys who score below the 30 percentile are extremely unlikely to be automotive repairmen. On the other hand, the boys who score above the 90 percentile have a much higher probability of being an automotive repairman even though this probability is only .171.

---

*See Section II for a complete description of these tests.

One out of four boys above the 99 percentile becomes an automotive repairman.

The large contribution to the accounted-for variance made by the automotive test, large in relation to the other predictor, would seem to indicate that the student's interests are important. An interest inventory could supplement the tests results when the latter identify individuals who might become automotive repairmen. The automotive industry could devise means of making careers in the field more attractive to high school youth, so as to attract and sustain their interest. When an assembly-line worker with little training can earn more than the average mechanic in a shop with training, the qualities that for many individuals makes being a mechanic attractive—relative freedom and independence, self-reliance, self-esteem—may not offset the more immediate financial reward as an inducement to the younger man. The stories told of the very high earnings of some mechanics obscure the fact that, if a mechanic is sick and cannot work, he does not earn.

An industry effort to upgrade the field of automotive repair and to mitigate its disadvantages as a career field could capitalize on an interest shown by ninth and tenth grade boys and induce them to select it as an occupational goal. Such an effort or industry could make the work of the school guidance counselor much easier, and eventually narrow the gap between production, maintenance, and consumer satisfaction.

VIII. SUMMARY AND RECOMMENDATIONS

During the decade of development that the Survey entailed, it passed through a sequence of stages each dominated by specific concerns. Within the initial period the test battery assumed its present form; most of the second period was devoted to preparing for the administration of the Survey nationally, and for the management of its actual administration; the third was a period of keeping abreast of the whereabouts of the participants; whereas the final stage included validating the battery and compiling the report.

From the outset, the basic task was the construction of an instrument that would identify potential automotive repairmen among high school youth at the ninth or tenth grade level. The tests in use at that time had limitations which the investigators preferred to avoid. One of the authors of this report prepared a set of job descriptions which he analyzed, while others surveyed the research literature, in both instances to determine what the test should measure. There proved to be a variety of factors that characterized the occupational behavior of employed automotive repairmen. The belief that one could measure these factors by means of objective procedures led to the construction of a battery of paper and pencil tests designed to measure the factors previously identified. This battery in its administered form consisted of an automotive and general vocabulary test; a reasoning test, a game-playing test; a chart-plotting test and a survey of object visualization test, all but the last of which the members of the research team constructed. The complete Survey also included a biographical inventory to measure the socio-economic and attitudinal factors believed to be related to entry into the trade of automotive repair. The tests consisted of items that were representative of the range of a special index of difficulty, and that also correlated biserially with the total scores. The reliabilities and intercorrelations that the test provided after a series of trial administrations suggested that the tests were appropriate for the task. This report contains a complete description of each test and the inventory.

Subsequently the staff made the requisite arrangements for administering the National Aptitude Survey tests to a randomly-selected, representative, national sample of ninth and tenth grade students in public high schools within the continental United States. The target population numbered ten thousand students. The strict requirements of the method of sample selection imposed limitations on the obtained size: some schools were unable to schedule the tests at all, while others declined to cooperate. In addition there were schools that had to defer the test period until the following semester. Salaried supervisors administered the tests under specified conditions. Each completed test booklet, upon receipt at the Survey office, received an identification number. A program especially prepared for The University of Michigan's computer converted the

raw data into test scores. A short period of time after completion of the testing, there was in storage a data card for each participant, which contained the test scores obtained by two methods of calculation. A tabulation of the data indicated that more than seven thousand male high school students in grades nine and ten had responded to the tests in the battery, (N=7265) and that there were complete scores for more than ninety-eight percent of these. A doctoral dissertation completed during this period reported reliability coefficients and test intercorrelations based on a representative sample of more than one thousand cases that are similar to those obtained in the pilot studies.

The intent to determine the predictive validity of the test battery made necessary a waiting period of six years until the participants could have accumulated sufficient post-high school experience. The emphasis in this period was mainly clerical. The problem of maintaining contact with more than seven thousand participants within the continental limits of the United States restricted the attention of the research team to meeting its challenge. There were three follow-ups during this period, each of which required a large number of mailings to the participants to keep their information on file current. So long as the majority of the sample population remained in high school the percentage of returned questionnaires hovered close to ninety-seven percent. After graduation returns diminished to approximately eighty percent.

The final stage of the research encompassed the identification of those participants who had actually become automotive repairmen and the determination of the battery's predictive validity. The research analyst formed two groups, one of which contained the scores of the automotive repairmen, and the other, the scores for one out of every seven men that remained. The test scores of this latter sample did not differ significantly from the scores of the remaining non-automotive repairmen population. The expectation was that certain scores obtained by automotive repairmen would differ significantly from those scores included in the representative sample. A stepwise linear regression program compared both sets of scores and indicated the relative contribution of the six test scores as well as every inventory item. The point at which the contribution of each factor relative to the others became negligible was the cutting point for the specification of the factors worthy of consideration. The multiple correlation coefficient at this point was .444 and the corresponding coefficient of multiple determination was .197. Two tests, the automotive and the general vocabulary and six biographical inventory items, were the best predictors. The two tests accounted for 14.5 percent of the variance in the respective participants' occupations. Through the use of the two test scores alone it was possible to develop a guide for potential users. The guide consists of a distribution of the percentages of automotive repairmen and others according to percentile ranges. It provides a means of estimating a ninth or tenth grade boy's chances of becoming an automotive repairman within three years after his expected graduation. This analysis also generated an equation, subsequently cross-validated, to use in computing the required score for determining a boy's chances.

APPENDIX A

NATIONAL APTITUDE TEST SCORES AS PREDICTORS OF ACHIEVEMENT IN HIGH SCHOOL VOCATIONAL EDUCATION COURSES

(Kaltsounis' Study)

TABLE 11

PREDICTORS AND CRITERIA (Kaltsounis' Study)

PREDICTORS	CRITERIA
"Number Right Score"	
1. Reasoning	I. Arts and Crafts
2. General Vocabulary	II. Electricity
3. Auto Vocabulary	III. Machine Shop
4. Chart Plotting	IV. Metals
5. Game Playing	V. Auto Shop
6. Survey Obj. Visualization	VI. Wood Shop
	VII. Drafting
"Formula Score"	VIII. Industrial Arts
	IX. Bookkeeping
7. Reasoning	X. Business Math
8. General Vocabulary	XI. General Business
9. Auto Vocabulary	XII. Typing
10. Chart Plotting	
11. Game Playing	
12. Survey Obj. Visualization	

TABLE 12

VALIDITY COEFFICIENTS OF THE NATIONAL APTITUDE SURVEY FOR TWELVE CRITERION GROUPS (Kaltsounis' Study)

	TEST[a]	CRITERION GROUP[a]											
		I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
NO. RIGHT SCORE	1	-.008	.417[b]	-.507	.204	.227	.110	.156[c]	.170[c]	.009	.310[c]	.188[b]	.277[c]
	2	.045	.130	-.667[b]	.177	.192	.151	.112[b]	.173[c]	.244[b]	.370[c]	.328[c]	.370[c]
	3	.064	.097	-.206	.338[c]	.393[b]	.156	.181[c]	.263[c]	.099	.074	.234[c]	.166[c]
	4	-.004	-.053	.685[b]	.304[c]	.256	.197[b]	.251[c]	.319[c]	.316[c]	.386[c]	.298[c]	.358[c]
	5	.159	.258	.148	.320[c]	.252	.212[c]	.287[c]	.273[c]	.153	.317[c]	.257[c]	.328[c]
	6	-.214	.483[c]	.350	.304[c]	.115	.353[c]	.448[c]	.312[c]	.001	.214[b]	.084	.285[c]
FORMULA SCORE	7	-.022	.410[b]	-.442	.208	.245	.117	.167[c]	.171[c]	.003	.308[c]	.074	.282[c]
	8	.041	.168	-.671[b]	.173	.162	.147	.103	.152[c]	.193	.345[c]	.437[c]	.361[c]
	9	.055	.148	.013	.343[c]	.372[b]	.153	.168[c]	.253[c]	.076	.045	.094	.138[c]
	10	-.026	-.047	.674[b]	.307[c]	.265	.207[b]	.236[c]	.314[c]	.310[c]	.399[c]	.333[c]	.355[c]
	11	.115	.291	.249	.295[c]	.244	.218[c]	.262[c]	.255[c]	.104	.294[c]	.261[c]	.298[c]
	12	-.268	.525[c]	.351	.303[c]	.111	.343[c]	.453[c]	.303[c]	.010	.208[b]	.124	.279[c]
	N	26	30	11	86	31	150	302	494	90	111	169	368

[a] Names of tests and criterion groups are found in Table 11.

[b] Significant at the .05 level.

[c] Significant at the .01 level.

TABLE 13

REGRESSION COEFFICIENTS, COEFFICIENTS OF CORRELATION AND DETERMINATION, AND STANDARD ERRORS OF ESTIMATE FOR THE CORRESPONDING CRITERION GROUPS (Kaltsounis' Study)

			CRITERION GROUP[a]										
			I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Constant[d]				.672		.565	.933	.995	.975	.678	.739	.029	.721	-.519
TEST	NO. RIGHT SCORE	1												
		2												
		3										.041	-.058	.029
		4					.060			.020				
		5							.021	.022	.050			.024
		6				.054				.016				.120
	FORMULA SCORE	7						.04		.020				
		8												.013
		9				.041							.098	
		10												
		11										.039		
		12		.057					.042				.026	-.073
r or R				.525[c]	.459[c]	.393[b]	.353[c]	.471[c]		.425[c]	.316[c]	.462[c]	.507[c]	.488[c]
r^2 or R^2				.2754	.2106	.1541	.1248	.2219		.1804	.1800	.2137	.2575	.2377
S.E. of Y				.895	.756	.990	.844	.888		.763	1.084	.952	.830	.799

[a] Names of tests and criterion groups are found in Table 11.

[b] Significant at the .05 level.

[c] Significant at the .01 level.

[d] Constant term ('a' coefficient) entering regression equation.

APPENDIX B

BIOGRAPHICAL INVENTORY ITEMS SELECTED BY THE REGRESSION PROGRAM

Biographical Inventory Items Selected by The Regression Program

Item	Item	Inventory Location
A	Which statement do you agree with? Ans. I think I would like the work of a garage mechanic.	p. 12, 10
B	What do you like to read most in magazines? Ans. Sports	p. 5, 4
C	Which would you rather do? Ans. Get a job when you finish high school	p. 11, 5
D	What do you use your own money for? Ans. Buying tools	p. 5, 3
E	Rank order the jobs you would like best. Ans. Machinist	p. 5, unnumbered
F	What kind of books do you like to read most? Ans. Lives of famous people	p. 5, 16

BIBLIOGRAPHY

1. American Personnel and Guidance Association. "The Use of Multifactor Tests in Guidance." A Reprint Series from The Personnel and Guidance Journal, September 1956-September 1957.

2. Automobile Manufacturers Association. Automobile Facts and Figures. Detroit: Automobile Manufacturers Association, 1963.

3. _____. Standards for Automotive Service Instruction in Secondary Schools. Automobile Manufacturers Association, 1965.

4. Bennett, G. K. "The DAT—A Seven-Year Follow-up." Test Service Bulletin No. 49, The Psychological Corporation, 1955.

5. _____ and Doppelt, J. E. "The Evaluation of Pairs of Tests for Guidance Use." Educational Psychological Measurement, 1948, 8, 319-325.

6. _____, Fry, D. E. and Owens, W. A. Test of Mechanical Comprehension. New York: Psychological Corporation, 1951.

7. _____, Seashore, H. G. and Wesman, A. G. "Validation of the Differential Aptitude Tests." First, Second, Third, Fourth, and Fifth Research Reports, The Psychological Corporation, 1948, 1949, 1949, 1951, 1951.

8. _____, Seashore, H. A. and Wesman, A. G. Differential Aptitude Tests. New York: Psychological Corporation, 1950.

9. _____, Seashore, H. G. and Wesman, A. G. "The Differential Aptitude Tests and Success in Vocational and Educational Fields." American Psychologist, 1951, 6, 296 (Abstract).

10. _____, Seashore, H. G. and Wesman, A. G. A Manual for the Differential Aptitude Tests. New York: The Psychological Corporation, 2d edition, 1952.

11. _____, Seashore, H. G. and Wesman, A. G. Supplementary Tables for the Differential Aptitude Tests. Washington: American Documentation Institute, 1952, No. 3656.

12. Berdie, R. F. "The Differential Aptitude Tests as Predictors in Engineering Training." Journal of Educational Psychology, 1951, 42, 114-123.

13. Bergsten, J. W. "A Nationwide Sample of Girls from School Lists." Journal of Experimental Education, 1958, 26, 197-208.

14. Bittner, R. H. and Wilder, C. E. "Expectancy Tables: A Method of Interpreting Correlation Coefficients." Journal of Experimental Education, 1946, 14, 245-252.

15. Bloom, B. S. and Peters, F. R. Academic Prediction Scales. Glencoe, Illinois: Free Press, 1961.

16. Brogden, H. J. "An Approach to the Problem of Differential Prediction." Psychometrika, 1946, 11, 139-154.

17. Cochran, W. G. Sampling Techniques. New York: Wiley, 1953.

18. Committee on Research and Publications. Definitions of Terms in Vocational and Practical Arts Education. Washington, D. C.,: American Vocational Association, 1954.

19. Cowden, D. J. "The Multiple-Partial Correlation Coefficient." Journal of the American Statistical Association, 1952, 47, 442-456.

20. Cronbach, L. J. Essentials of Psychological Testing. 2d. edition. New York: Harper, 1960.

21. Dailey, J. T. and Shaycoft, M. F. Types of Tests in Project Talent. Washington: United States Government Printing Office, 1961.

22. Dallemand, J. E. "Stepwise Regression Program on the IBM 704." Warren, Michigan: Research Laboratories, General Motors Corporation, 1958 (GMR-199).

23. _____, "Stepwise Regression—1962 Supplement." Warren, Michigan: Research Laboratories, General Motors Corporation, 1962.

24. Doppelt, J. E., Seashore, H. G., and Odgers, J. G. "Validation of the Differential Aptitude Tests for Auto Mechanics and Machine Shop Students." Unpublished Report, The Psychological Corporation.

25. Dvorak, B. J. "New G.A.T.B. Occupational Aptitude Pattern Norm Structure." Vocational Guidance Quarterly, 1955, 3, 110-112.

26. Ebel, R. L. "Obtaining and Reporting Evidence on Content Validity." Educational Psychological Measurement, 1956, 16, 269-282.

27. Fink, D. C. "The Efficiency of Certain Criteria in Predicting School Dropout." Dissertation Abstracts, 1962, 1555.

28. Fisher, R. A. Statistical Methods for Research Workers. 11th edition. New York: Hafner, 1950.

29. Flanagan, J. E. "The Effectiveness of Short Methods for Calculating Correlation Coefficients." Psychological Bulletin, 1952, 49, 342-348.

30. _____. Aptitude Classification Tests. 14-test Edition. Chicago: Science Research Associates, 1953.

31. _____. Technical Supplement. Chicago: Science Research Associates, 1954.

32. _____. Interpreting Test Scores. Chicago: Science Research Associates, 1956.

33. _____. Technical Report—Flanagan Aptitude Classification Tests. Chicago: Science Research Associates, 1959.

34. _____. "Project TALENT: The First National Census of Aptitudes and Abilities." National Council on Measurements Used in Education, 1960, 17, 37-44.

35. _____, and Daily, J. T. "Project TALENT: The Identification, Development and Utilization of Human Talents." Personnel Guidance Journal, 1960, 38, 504-505.

36. Fowler, F. M. Selection of Students for Vocational Training. Washington, D.C.: United States Office of Education, 1945.

37. General Aptitude Test Battery—III: Development. Washington, D. C.: United States Department of Labor, 1962.

38. Ghiselli, E. E. Measurement of Occupational Aptitude. Berkeley: University of California Press, 1955.

39. _____. "The Measurement of Occupational Aptitude." University of California Publications in Psychology, 1955, 8 (2), 101-216.

40. _____, and Brown, J. K. "Validity of Tests for Auto Mechanics." Journal of Applied Psychology, 1951, 35, 23-24.

41. Goodenough, F. L. Mental Testing, Its History, Principles, and Applications. New York: Rinehart and Co., Inc., 1949.

42. Goslin, D. A. The Search for Ability. New York: Russell Sage Foundation, 1963.

43. Green, R. F., Guilford, J. P., Christensen, P. R., and Comrey, A. L. "A Factor-Analytic Study of Reasoning Abilities." Psychometrika, 1953, 18, 135-160.

44. Guilford, J. P. "Factor Analysis in a Test-Development Program." Psychological Review, 1948, 55, 79-94.

45. _____. Fundamental Statistics in Psychology and Education. 3rd edition. New York: McGraw-Hill Book Co., 1956.

46. _____, Berger, R. M., and Christensen, P. R. "A Factor-Analytic Study of Planning: I. Hypotheses and Description of Tests." Reports from the Psychological Laboratory, No. 10. University of Southern California, 1954.

47. _____, Berger, R. M., and Christensen, P. R. "A Factor-Analytic Study of Planning: II. Administration of Tests and Analysis of Results." Reports from the Psychological Laboratory, No. 12. Los Angeles: University of Southern California, 1955.

48. _____, Christensen, P.R., Kettner, N. W., Green, R. F., and Hertzka, A. F. "A Factor-Analytic Study of Navy Reasoning Tests with the Air Force Aircrew Classification Battery." Educational Psychological Measurement, 1954, 14, 301-325.

49. _____, Kettner, N. W., and Christensen, P. R. "The Nature of the General Reasoning Factor." Psychological Review, 1956, 63, 169-172.

50. _____, and Zimmerman, W. S. The Guilford-Zimmerman Aptitude Survey, Manual of Instructions and Interpretations. 2d edition. Beverly Hills, California: Sheridan Supply Company, 1956.

51. Guion, R. M., Personnel Testing. New York: McGraw-Hill, 1965.

52. Gulliksen, H. Theory of Mental Tests. New York: John Wiley and Sons, 1950.

53. Hansen, M. H. "Modern Methods in the Sampling of Human Populations." American Journal of Public Health, 1951, 41, 662-668.

54. Harper, A. E., Jr. "How to Develop Local Norms." Journal of Vocational Educational Guidance, India, 1959, 5, 106-119.

55. Harrell, T. W. "A Factor Analysis of Mechanical Ability Tests." Psychometrika, 1940, 5, 17-33.

56. Hensel, J. W. "Relation of High School Course Work to Achievement at the Iowa State University of Science and Technology." Dissertation Abstracts, 1962, 885.

57. Holzinger-Crowder Uni-Factor Tests, Harcourt, Brace and World, Inc.

58. Horst, A. P. "Regression Weights as a Function of Test Length." Psychometrika, 1948, 13, 125-134.

59. _____. "Determination of Optimal Test Length to Maximize the Multiple Correlation." Psychometrika, 1949, 14, 79-88.

60. Ibrahim, Y. A. "A Study of the Effects of Item Analysis Upon the Predictive Efficiency of a Standardized Test." Dissertation Abstracts, 1962, 22, 2283.

61. Jenkins, J. G. "Validity for What?" Journal of Consulting Psychology, 1946, 10, 93.

62. Joint Committee of the American Psychological Association. "Technical Recommendations for Psychological Tests and Diagnostic Techniques." Supplement to the Psychological Bulletin, March 1954, 51, (2), Part 2, 201-238.

63. Kaltsounis, G. L. "National Aptitude Survey Test Scores as Predictors of Achievement in High School Vocational Education Courses." Unpublished doctoral dissertation, The University of Michigan, 1965.

64. _____. "Summary of the Tests, the Tested Sample, the Meaning of Scores." University of Michigan-Automobile Manufacturers Association, 1962 (Unpublished report).

65. Katona, G., Kish, L., Lansing, J. B., and Dent, J. K. "Methods of the Survey of Consumer Finances." Federal Reserve Bulletin, 1950, 36, No. 2, 795-809.

66. Kirkpatrick, E. A. "A Vocabulary Test." Popular Science Monthly, 1907, 70, 157-164.

67. Kish, L. "A Two-stage Sample of a City." American Sociological Review, 1952, 17, 761-769.

68. _____. "Selection of the Sample." In Festinger, L. and Katz, D. Research Methods in the Behavioral Sciences. New York: The Dryden Press, 1953, 175-239.

69. _____. "Confidence Intervals for Clustered Samples." American Sociological Review, 1957, 22, 154-165.

70. Littleton, I. T. "Prediction in Auto Trade Courses." Journal of Applied Psychology, 1952, 36, 15-19.

71. Lord, F. M. "Test Norms and Sampling Theory." Journal of Experimental Psychology, 1959, 27, 247-263.

72. Manual for the Multiple Aptitude Tests. Los Angeles: California Test Bureau, 1955.

73. Masten, S. H. "The Value of Entrance Tests and Nine High School Variables in the Selection of Freshmen at Hofstra College: A Study in Differential Prediction." Dissertation Abstracts, 1962, 23, 886-887.

74. Michael, W. B., Zimmerman, W. S., and Guilford, J. P. "An Investigation of the Nature of Spatial-Relations and Visualization Factors in Two High School Samples." Educational Psychological Measurement, 1951, 11, 561-567.

75. Mitchell, B. C. "The Relation of High School Achievement to the Abilities Measured by the Holzinger-Crowder Uni-Factor Tests." Educational Psychological Measurement, 1955, 15, 487-490.

76. Mollenkopf, W. G. "Time Limits and the Behavior of Test Takers." Educational Psychological Measurement, 1960, 20, 223-230.

77. Mosier, C. I., Cureton, E. E., Katzell, R. A., and Wherry, R. J. "The Need and Means of Cross-Validation." Educational and Psychological Measurement, 1951, 11.

78. Nerden, J. T: "Industrial Education in a Modern Industrial Economy." Address at Ohio Vocational Association Convention, Columbus, Ohio: September 27, 1958.

79. Nicholson, T. "Who'll Keep Cars Going?" Detroit Free Press, August 15, 1957.

80. Novak, B. J., Scheuhing, M. A. "Predicting Success in High School Industrial Courses." Industrial Arts and Vocational Education, 1951, 40, 391-394.

81. O'Rourke, L. J. O'Rourke Mechanical Aptitude Test. Psychological Institute, 1939.

82. Paterson, D. G., Elliott, R. M., Anderson, L. D., Toops, H. A., and Heidbreder, Edna. The Minnesota Mechanical Abilities Test. Minneapolis: University of Minnesota Press, 1930.

83. Pilliner, A. E. G. "The Application of Analysis of Variance to Problems of Correlation." British Journal of Psychology, Statistical Section, 1952, 5, 31-38.

84. Potter, R. E. "Testing for Mechanical Occupations." Vocational Guidance Quarterly, 1953, 1, 12-14.

85. Rothney, J. W. M. "Interpreting Test Scores to Counselees." Occupations, 1952, 30, 320-322.

86. Seashore, H. G. "Tenth Grade Tests as Predictors of Twelfth Grade Scholarship and College Entrance Status." Journal of Counseling Psychology, 1954, 1, 106-115.

87. _____. "Cross-Validation of Equations for Predicting CEEB-SAT Scores from DAT Scores." Journal of Counseling Psychology, 1955, 2, 229-230.

88. Segel, D. and Raskin, E. Technical Report on the Multiple Aptitude Tests. Monterey, California: California Test Bureau, 1959 Edition.

89. Senders, V. L. Measurement and Statistics. New York: Oxford University Press, 1958.

90. Smith, J. A. "Studies in Industrial Education: An Annotated Bibliography of Theses and Essays." Wayne State University, 1957.

91. Stenquist, J. L. Mechanical Aptitude Tests. Yonkers, New York: World Book Company, 1922.

92. Stoughton, R. W. "The Differential Predictive Values of the Differential Aptitude Tests in the Connecticut Technical Schools." Unpublished doctoral dissertation, University of Connecticut, 1955.

93. Summerfield, A. and Lubin, A. "A Square Root Method of Selecting a Minimum Set of Variables in Multiple Regressions." Psychometrika, 1951, 16, 271-284, 425-437.

94. Super, D. E. Appraising Vocational Fitness by Means of Psychological Tests. New York: Harper and Brothers, 1949, 221-281.

95. Thorndike, R. L. Personnel Selection, Test and Measurement Techniques. New York: John Wiley and Sons, Inc., 1949.

96. _____ and Hagen, E. P. Measurement and Evaluation in Psychology and Education. New York: Wiley, 1961.

97. _____ and Hagen, E. P. "Validation of the Electronics Techician Selection Test at Selected Class 'A' Schools." Bureau of Naval Personnel, Technical Bulletin, 1955, 55.

98. Thurstone, L. L. "Primary Mental Abilities." Psychometrika, 1938, 1.

99. Turse, P. L. Clerical Aptitude Tests. New York: World Book Company, 1955.

100. United States Employment Service, Department of Labor. "Training and Reference Manual for Job Analyses." June 1944.

101. United States Employment Service. General Aptitude Test Battery. Washington, D. C.: United States Employment Service, 1947.

102. University of Pittsburgh, Project Talent Office. "The Story of Project Talent." Bulletin No. 1, 1959.

103. Walter, H. M. and Lev, J. Statistical Inference. New York: Henry Holt and Company, 1953.

104. Walston, E. B. "The Autobiography in the Prediction of College Field of Concentration." Unpublished Ph.D. dissertation, Harvard University, 1958.

105. Wesman, A. G. "Expectancy Tables—A Way of Interpreting Test Validity." Test Service Bulletin, No. 38. The Psychological Corporation, 1949.

106. _____. "The Differential Aptitude Tests—A Five Year Report." Personnel Guidance Journal, 1942, 31, 167-175.

107. _____ and Bennett, G. K. "Problems of Differential Prediction." Educational Psychological Measurement, 1951, 11, 265-272.

108. West, V. I. "Replacing Variables in Correlation Problems." Journal of the American Statistical Association, 1952, 47, 185-190.

109. Westervelt, F. H. "A Study of Automatic System Simulation Programming and the Analysis of the Behavior of Physical Systems Using an Internally Stored Program Computer." Unpublished Ph.D. dissertation, University of Michigan, 1960. (Microfilm available at University Microfilms, Ann Arbor, Michigan: Library of Congress No. 61, 1807.)

110. _____. "Automatic System Simulation Programming." (Reprinted from doctoral thesis.) Industry Program of the College of Engineering, The University of Michigan, November, 1960.

111. _____. "Analysis and Synthesis of Engineering Systems by Digital-Computer Programming." The American Society of Mechanical Engineers, No. 62-WA-214, 1962.

112. _____. "Regression Analysis Program on the 709/7090." Computing Center, The University of Michigan. (Copies of write-up available at the center.)

113. Williams, N. "A Study of the Validity of the Verbal Reasoning Subtest and the Abstract Reasoning Subtest of the Differential Aptitude Tests." Educational Psychological Measurement, 1952, 12, 129-131.

114. Wolbers, H. L., Jr. "Estimation of Correlational Relationships Using the Technique of Sequential Analysis." American Psychologist, 1950, 5, 463.